L^p ESTIMATES FOR A SINGULAR INTEGRAL OPERATOR MOTIVATED BY CALDERÓN'S SECOND COMMUTATOR

EYVINDUR ARI PALSSON

ABSTRACT. We prove a wide range of L^p estimates for a trilinear singular integral operator motivated by dropping one average in Calderón's second commutator. For comparison by dropping two averages in Calderón's second commutator one faces the trilinear Hilbert transform. The novelty in this paper is that in order to avoid difficulty of the level of the trilinear Hilbert transform, we choose to view the symbol of the operator as a non-standard symbol. The methods used come from time-frequency analysis but must be adapted to the fact that our symbol is non-standard.

1. INTRODUCTION

1.1. **History.** The k-th Calderón commutator, $k \in \{1, 2, 3, \ldots\}$, is given by

$$\mathcal{C}_A^{(k)} f(x) = p.v. \int_{\mathbb{R}} \frac{1}{x-y} \left(\frac{A(x) - A(y)}{x - y} \right)^k f(y) dy$$

where A is Lipschitz and $A' \in L^\infty(\mathbb{R})$. Calderón studied these operators in connection with an algebra of pseudo-differential operators. He was also motivated by possible applications to operators of the type

$$(1.1) \qquad p.v. \int_{\mathbb{R}} \frac{1}{x-y} \ F\left(\frac{A(x) - A(y)}{x - y} \right) f(y) \ dy$$

where F is an analytic function. The Cauchy integral on Lipschitz curves and double layer potentials are examples of the previous operator. In 1965 Calderón showed

$$\mathcal{C}_A^{(k)} : L^p \to L^p \text{ for } 1 < p < \infty$$

for $k = 1$ [2]. Coifman and Meyer extended his result in 1975 to $k = 2, 3, \ldots$ [4]. The estimates obtained did not clearly indicate how the boundedness constant depended on k. Building on the work of Coifman and Meyer, Calderón was able to prove the above estimates with a boundedness constant that depended on k exponentially. This way he was able to prove bounds for operators of the type (1.1), as long as the Lipschitz constant was small. Finally, in 1982 Coifman, McIntosh and Meyer showed the above estimates with a boundedness constant that depended on k polynomially [5] and were thus able to show a wide range of L^p estimates for operators of the type (1.1).

1.2. **Motivation.** Calderón observed that one can write the following as an average

$$\frac{A(x) - A(y)}{x - y} = \int_0^1 A'(x + \alpha(y - x)) d\alpha.$$

Using this trick and a substitution he rewrote his first commutator as

$$\mathcal{C}_A^{(1)} f(x) = \int_0^1 \int_{\mathbb{R}} A'(x + \alpha t) f(x + t) \frac{1}{t} dt d\alpha.$$

He then asked if one dropped the average and fixed α whether L^p estimates could be found for the resulting operator, uniformly in α. This motivated the definition of the bilinear Hilbert transform

$$BHT_\alpha(f_1, f_2)(x) = p.v. \int_{\mathbb{R}} f_1(x + \alpha t) f_2(x + t) \frac{1}{t} dt.$$

In two papers from 1997 and 1999, Lacey and Thiele showed that the bilinear Hilbert transform BHT_α maps $L^p \times L^q$ into L^r when $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, $1 < p, q \le \infty$ and $\frac{2}{3} < r < \infty$ with a bound depending on α [9, 10]. Uniform boundedness of these L^p estimates was resolved later [7, 17]. Note that r only goes down to $\frac{2}{3}$, not $\frac{1}{2}$ as one would expect from Hölder type estimates. It is still an open problem whether r can be pushed all the way down to $\frac{1}{2}$.

In a similar fashion then one can rewrite the second Calderón commutator with two averages. Dropping both averages motivates the definition of the trilinear Hilbert transform.

$$THT_{\vec{\alpha}}(f_1, f_2, f_3)(x) = p.v. \int_{\mathbb{R}} f_1(x + \alpha_1 t) f_2(x + \alpha_2 t) f_3(x + t) \frac{1}{t} dt$$

In contrast to the bilinear Hilbert transform then no L^p estimates are known for the trilinear Hilbert transform.

In this paper we will study a trilinear operator motivated by $C_A^{(2)}$ in a similar fashion as $THT_{\vec{\alpha}}$, except we drop one average, not two. Define

$$(1.2) \qquad T_\beta(f_1, f_2, f_3)(x) := p.v. \int_{\mathbb{R}} \left(\int_0^1 f_1(x + \alpha t) d\alpha \right) f_2(x + \beta t) f_3(x + t) \frac{1}{t} dt.$$

1.3. **Known estimates.** Benyi, Demeter, Nahmod, Thiele, Torres and Villarroya obtained a modulation invariant bilinear $T(1)$ theorem [1]. If one fixes $f_1 \in L^\infty(\mathbb{R})$ and looks at the bilinear operator

$$(f_2, f_3) \mapsto p.v. \int_{\mathbb{R}} \frac{\left(\int_0^1 f_1(x + \alpha t) d\alpha \right)}{t} f_2(x + \beta t) f_3(x + t) dt,$$

one can apply their theorem to obtain the following L^p estimates for T_β

$$T_\beta : L^\infty \times L^{p_1} \times L^{p_2} \to L^p$$

for $\beta \notin \{0, 1\}$ if $\dfrac{1}{p_1} + \dfrac{1}{p_2} = \dfrac{1}{p}$, $1 < p_1, p_2 \le \infty$ and $\frac{2}{3} < p < \infty$. These are the only known L^p estimates for T_β.

1.4. **Result.** The main theorem of this paper establishes the following wide range of L^p estimates for T_β.

Theorem 1.1. *Let* $\beta \notin \{0, 1\}$, $1 < p_1, p_2, p_3 \le \infty$,

$$\frac{1}{2} < p := \frac{p_1 p_2 p_3}{p_1 p_2 + p_1 p_3 + p_2 p_3} < \infty \quad and \quad \frac{2}{3} < \frac{p_2 p_3}{p_2 + p_3} \le \infty.$$

Then there exists a constant C_{β,p_1,p_2,p_3} such that

$$\|T_\beta(f_1, f_2, f_3)\|_p \leq C_{\beta,p_1,p_2,p_3}\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}$$

for all f_1, f_2 and f_3 in $\mathcal{S}(\mathbb{R})$.

The theorem recovers all known L^p estimates for the operator. Known L^p estimates for both the bilinear Hilbert transform and for Calderón's first commutator follow as a corollary.

Compared to the theorem on the bilinear Hilbert transform, this theorem has an extra condition.

$$\frac{2}{3} < \frac{p_2 p_3}{p_2 + p_3} \leq \infty$$

This condition implies that we have not improved the previously known L^p estimates for the bilinear Hilbert transform. We also require the condition $\frac{1}{2} < p$, which is not the largest possible range of L^p estimates expected. Based on the known estimates for the bilinear Hilbert transform one would expect to be able to go all the way down to $\frac{2}{5}$. This remains an open problem.

Note that if $\beta = 0, 1$ then we obtain trilinear operators that only involve multiplication of functions and the first Calderón commutator. The L^p-bounds of these operators are easy to determine.

1.5. **Approach.** The standard way of understanding the boundedness of the Calderón commutators is to use the $T(1)$ theorem. In order to use such an approach on T_β we would need some sort of a trilinear $T(1)$ theorem. Despite the existence of some multilinear $T(1)$ theorems [3, 8] then there is no such appropriate theorem for T_β. The other canonical way of trying to understand T_β would be to establish uniform L^p estimates on the trilinear Hilbert transform. Since no L^p estimates exist, uniform estimates are out of reach. The obvious approaches to find L^p estimates fail so we need some novel ideas.

On the Fourier side it is equivalent to show L^p estimates for an operator T_β given by

$$(1.3) \quad T_\beta(f_1, f_2, f_3)(x) = \int_{\mathbb{R}^3} \left[\int_0^1 \mathsf{sgn}(\alpha\xi_1 + \beta\xi_2 + \xi_3)d\alpha \right] \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)$$

$$e^{2\pi i(\xi_1 + \xi_2 + \xi_3)x} d\xi_1 d\xi_2 d\xi_3.$$

where sgn is the usual sign function. The symbol $\int_0^1 \mathsf{sgn}(\alpha\xi_1 + \beta\xi_2 + \xi_3)d\alpha$ has a singularity around the line $\xi_1 = 0$, $\beta\xi_2 + \xi_3 = 0$ in the sense that it is discontinuous. This is similar to the bilinear Hilbert transform. Unlike standard symbols, which are assumed to be smooth outside the set where they are singular, this symbol is continuous but not differentiable on the planes $\xi_1 + \beta\xi_2 + \xi_3 = 0$ and $\xi_1 = 0$ away from the previous line. We approach the symbol as a rough non-standard symbol and use techniques in the spirit of the bilinear Hilbert transform. An important ingredient in that approach are new proofs of the L^p estimates for the Calderón commutators by Muscalu [11]. The techniques and notation are also heavily inspired by Muscalu, Tao and Thiele [12, 13].

There exist theorems that give immediate L^p estimates for operators with standard symbols where the dimension of the singularity is strictly less than half the dimension of the frequency space of the form associated to the operator [15]. Even if our symbol had been standard outside the line then those kind of theorems would not have been applicable because the line is degenerate.

1.6. **Acknowledgements.** The author would like to thank his thesis adviser, Camil Muscalu, for his guidance and many helpful conversations about this problem.

2. Notation

We use $A \lesssim B$ to denote the statement that $A \leq CB$ for some large constant C and $A \ll B$ to denote the statement that $A \leq C^{-1}B$ for some large constant C. Our constants C shall always be independent of the tiles $\vec{P}$.

Given any interval I, let $|I|$ denote the Lebesgue measure of I and let cI denote the interval with the same center as I but c times the side-length. Also define the approximate cutoff function $\tilde{\chi}_I$ by

$$\tilde{\chi}_I(x) := (1 + (\frac{|x - x_I|}{|I|})^2)^{-1/2}$$

where x_I is the center of I.

Define $\langle n \rangle := 2 + |n|$ for $n \in \mathbb{Z}$.

3. Symbol

The meaning of (1.2) is

(3.1)
$$\lim_{\epsilon \to 0^+} \int_{|t| > \epsilon} \left(\int_0^1 f_1(x + \alpha t)d\alpha \right) f_2(x + \beta t)f_3(x + t)\frac{1}{t}dt$$

where the limit exists. Assume f_1, f_2 and f_3 are Schwartz functions on $\mathbb{R}$. We will show that (3.1) exists in that case and we will rewrite it in a convenient way.

Write (3.1) as

$$\lim_{\substack{\epsilon \to 0^+ \\ N \to \infty}} \int_{\epsilon < |t| < N} \left[\int_0^1 \int_{\mathbb{R}} \widehat{f_1}(\xi_1)e^{2\pi i \xi_1(x + \alpha t)}d\xi_1 \ d\alpha \right]$$

$$\int_{\mathbb{R}} \widehat{f_2}(\xi_2)e^{2\pi i \xi_2(x + \beta t)}d\xi_2 \int_{\mathbb{R}} \widehat{f_3}(\xi_3)e^{2\pi i \xi_3(x + t)}d\xi_3 \ \frac{1}{t}dt$$

which is equal to

$$\lim_{\substack{\epsilon \to 0^+ \\ N \to \infty}} \int_{\epsilon < |t| < N} \int_{\mathbb{R}^3} \left[\int_0^1 \frac{1}{t}e^{-2\pi i(-\alpha\xi_1 - \beta\xi_2 - \xi_3)}d\alpha \right] \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)e^{2\pi i x(\xi_1 + \xi_2 + \xi_3)}d\xi_1 d\xi_2 d\xi_3 dt$$

The function being integrated, viewed as depending on ξ_1, ξ_2, ξ_3 and t is clearly absolutely integrable on $\mathbb{R}^4$ and by applying Fubini's theorem together with dominated convergence we see that the formula becomes equivalent to

(3.2)
$$\int_{\mathbb{R}^3} \left[\int_0^1 \mathsf{sgn}(-\alpha\xi_1 - \beta\xi_2 - \xi_3)d\alpha \right] \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)e^{2\pi i(\xi_1 + \xi_2 + \xi_3)x}d\xi_1 d\xi_2 d\xi_3$$

which clearly exists since $\widehat{f_1}$, $\widehat{f_2}$ and $\widehat{f_3}$ are also Schwartz functions.

A product of three functions satisfies a Hölder type inequality as we obtain in Theorem 1.1. Since the product can be written as

(3.3)
$$\int_{\mathbb{R}^3} \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)e^{2\pi i(\xi_1 + \xi_2 + \xi_3)x}d\xi_1 d\xi_2 d\xi_3$$

and using $\mathsf{sgn}(-x) = -\mathsf{sgn}(x)$ it becomes clear by subtracting (3.2) from (3.3) that it is enough to consider L^p estimates for

$$(3.4) \quad \tilde{T}_\beta(f_1, f_2, f_3)(x) := \int_{\mathbb{R}^3} \left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1 + \beta\xi_2 + \xi_3)d\alpha \right] \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)$$

$$e^{2\pi i(\xi_1 + \xi_2 + \xi_3)x} d\xi_1 d\xi_2 d\xi_3.$$

where $1_{\mathbb{R}_+}$ is the characteristic function for the positive real axis.

Similar to what was mentioned in the introduction then the symbol

$$\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1 + \beta\xi_2 + \xi_3)d\alpha$$

is not continuous around the line $\xi_1 = 0$, $\beta\xi_2 + \xi_3 = 0$, continuous but not differentiable around the planes $\xi_1 + \beta\xi_2 + \xi_3 = 0$ and $\beta\xi_2 + \xi_3 = 0$, away from the previous line, but smooth everywhere else. It is tempting to view the symbol as a trilinear symbol of the variables ξ_1, ξ_2, ξ_3. That would however result in a problem of the same difficulty as the trilinear Hilbert transform. We choose thus instead to view it as a non-standard bilinear symbol of the variables ξ_1 and $\beta\xi_2 + \xi_3$.

4. DISCRETIZATION

We will now come up with a "discretized" variant of the "continuous" form associated to (3.4). We start by reviewing some standard definitions and comments [13].

Definition 4.1. *Let $n \geq 1$ and $\sigma \in \{0, \frac{1}{3}, \frac{2}{3}\}^n$. We define the shifted n-dyadic mesh $D = D_\sigma^n$ to be the collection of cubes of the form*

$$D_\sigma^n := \{2^j(k + (0,1)^n + (-1)^j\sigma) | j \in \mathbb{Z}, \quad k \in \mathbb{Z}^n\}$$

We define a shifted dyadic cube to be any member of a shifted n-dyadic mesh.

Observe that for every cube Q, there exists a shifted dyadic cube Q' such that $Q \subseteq \frac{7}{10}Q'$ and $|Q'| \sim |Q|$; this is best seen by first verifying the $n = 1$ case.

Definition 4.2. *A subset D' of a shifted n-dyadic grid D is called sparse, if for any two cubes Q, Q' in D with $Q \neq Q'$ we have $|Q| < |Q'|$ implies $|10^9 Q| < |Q'|$ and $|Q| = |Q'|$ implies $10^9 Q \cap 10^9 Q' = \emptyset$.*

Observe that any subset of a shifted n-dyadic grid (with $n \leq 4$ say), can be split into $O(1)$ sparse subsets.

Definition 4.3. *Let $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4) \in \{0, \frac{1}{3}, \frac{2}{3}\}^4$, and let $1 \leq i \leq 4$. An i-tile with shift σ_i is a rectangle $P = I_P \times \omega_P$ with area 1 and with $I_P \in D_0^1$, $\omega_P \in D_{\sigma_i}^1$. A quadtile with shift σ is a 4-tuble $\vec{P} = (P_1, P_2, P_3, P_4)$ such that each P_i is an i-tile with shift σ_i, and the $I_{P_i} = I_{\vec{P}}$ are independent of i. The frequency cube $Q_{\vec{P}}$ of a quadtile is defined to be $\Pi_{i=1}^4 \omega_{P_i}$*

We sometimes refer to i-tiles with shift σ just as i-tiles, or even as tiles, if the parameters σ, i are unimportant.

Definition 4.4. *A set $\vec{\mathbf{P}}$ of quadtiles is called sparse, if all quadtiles in $\vec{\mathbf{P}}$ have the same shift and the set $\{Q_{\vec{P}} : \vec{P} \in \vec{\mathbf{P}}\}$ is sparse.*

Again, any set of quadtiles can be split into $O(1)$ sparse subsets.

Definition 4.5. *Let P and P' be tiles. We write $P' < P$ if $I_{P'} \subsetneq I_P$ and $5\omega_P \subseteq 5\omega_{P'}$, and $P' \leq P$ if $P' < P$ or $P' = P$. We write $P' \lesssim P$ if $I_{P'} \subseteq I_P$ and $10^7\omega_P \subseteq 10^7\omega_{P'}$. We write $P' \lesssim' P$ if $P' \lesssim P$ and $P' \not\leq P$.*

This ordering by Muscalu, Tao and Thiele [13] is in the spirit of that in Fefferman [6] or Lacey and Thiele [9, 10]. The main difference from the previous orderings is that P' and P do not quite have to intersect which turns out to be convenient for technical purposes.

Definition 4.6. *Let P be a tile. An L^p normalized wave packet on P, $1 \leq p < \infty$, is a function ϕ_P which has Fourier support in $\frac{9}{10}\omega_P$ and obeys the estimates*

$$|\phi_P(x)| \lesssim |I_P|^{-1/p}\tilde{\chi}_I(x)^M$$

for all $M > 0$, with the implicit constant depending on M.

Heuristically, ϕ_P is L^p-normalized and is supported in P.

Now that we have the tools from Muscalu, Tao and Thiele [13] then let us start decomposing. We start with two standard Littlewood-Paley decompositions and write

$$1_{\mathbb{R}}(\xi_1) = \sum_{k_1} \widehat{\Psi_{k_1}}(\xi_1)$$

and

$$1_{\mathbb{R}}(\beta\xi_2 + \xi_3) = \sum_{k_2} \widehat{\Psi_{k_2}}(\beta\xi_2 + \xi_3)$$

where as usual, $\widehat{\Psi_{k_1}}(\xi_1)$ and $\widehat{\Psi_{k_2}}(\beta\xi_2 + \xi_3)$ are bumps supported in the regions $|\xi_1| \sim 2^{k_1}$ and $|\beta\xi_2 + \xi_3| \sim 2^{k_2}$ respectively. In particular we get

$$(4.1) \qquad 1_{\mathbb{R}}(\xi_1, \beta\xi_2 + \xi_3) = \sum_{k_1, k_2} \widehat{\Psi_{k_1}}(\xi_1)\widehat{\Psi_{k_2}}(\beta\xi_2 + \xi_3)$$

By splitting (4.1) over the regions where $k_1 \ll k_2$, $k_2 \ll k_1$ and $k_1 \sim k_2$ we obtain the decomposition

$$(4.2) \qquad 1_{\mathbb{R}}(\xi_1, \beta\xi_2 + \xi_3) = \sum_k \widehat{\Psi_k}(\xi_1)\widehat{\Phi_k}(\beta\xi_2 + \xi_3) \quad +$$

$$(4.3) \qquad \sum_k \widehat{\Phi_k}(\xi_1)\widehat{\Psi_k}(\beta\xi_2 + \xi_3) \quad +$$

$$(4.4) \qquad \sum_{k_1 \sim k_2} \widehat{\Psi_{k_1}}(\xi_1)\widehat{\Psi_{k_2}}(\beta\xi_2 + \xi_3).$$

where $\widehat{\Phi_k}$ is a bump supported on an interval, symmetric with respect to the origin of length $\sim 2^k$.

Note that $\widehat{\Phi_k}(\beta\xi_2 + \xi_3)$ is supported in $\mathbb{R}^2$ on a strip around the line $\beta\xi_2 + \xi_3 = 0$ of width $\sim 2^k$. We can cover that strip with shifted dyadic cubes with side-length $\sim 2^k$. Similarly then $\widehat{\Psi_k}(\beta\xi_2 + \xi_3)$ is supported in $\mathbb{R}^2$ on two strips of width $\sim 2^k$ but this time away from $\beta\xi_2 + \xi_3 = 0$. Again we can cover those strips with shifted dyadic cubes of a similar scale.

Thus we come up with a decomposition

$$(4.5) \qquad a(\xi_1, \xi_2, \xi_3) = \sum_{\vec{Q} \in \vec{\mathbf{Q}}} \widehat{\phi_{Q_1,1}}(\xi_1) \widehat{\phi_{Q_2,2}}(\xi_2) \widehat{\phi_{Q_3,3}}(\xi_3)$$

for each of the three cases (4.2), (4.3), (4.4) such that

$$\frac{1}{10} < a(\xi_1, \xi_2, \xi_3) < 10.$$

Here $\phi_{Q_i,i}$ is an L^1 normalized wave packet on a tile $I_{\vec{Q}} \times Q_i$ for $i = 1, 2, 3$, where Q_i is a shifted dyadic interval that depends on the decomposition in each of the three cases and $I_{\vec{Q}}$ is a dyadic interval such that $|I_{\vec{Q}}| \sim |Q_i|^{-1}$ for $i = 1, 2, 3$.

Since $\xi_1 \in \frac{9}{10}Q_1$, $\xi_2 \in \frac{9}{10}Q_2$ and $\xi_3 \in \frac{9}{10}Q_3$ it follows that $\xi_1 + \xi_2 + \xi_3 \in \frac{9}{10}Q_1 + \frac{9}{10}Q_2 + \frac{9}{10}Q_3$ and as a consequence one can find a shifted dyadic interval Q_4 with the property that $\frac{9}{10}Q_1 + \frac{9}{10}Q_2 + \frac{9}{10}Q_3 \subseteq -\frac{7}{10}Q_4$ and also satisfying $|Q_1| = |Q_2| = |Q_3| \sim |Q_4|$. In particular there exists an L^1 normalized wave packet $\phi_{Q_4,4}$ adapted to $I \times Q_4$ such that $\widehat{\phi_{Q_4,4}} \equiv 1$ on $-\frac{9}{10}Q_1 - \frac{9}{10}Q_2 - \frac{9}{10}Q_3$.

Thus (4.5) can be written as

$$(4.6) \qquad a(\xi_1, \xi_2, \xi_3) = \sum_{\vec{Q} \in \vec{\mathbf{Q}}} \widehat{\phi_{Q_1,1}}(\xi_1) \widehat{\phi_{Q_2,2}}(\xi_2) \widehat{\phi_{Q_3,3}}(\xi_3) \widehat{\phi_{Q_4,4}}(-\xi_1 - \xi_2 - \xi_3)$$

where this time $\vec{\mathbf{Q}}$ is a collection of shifted dyadic quasi-cubes in $\mathbb{R}^4$. Modulo a finite refinement we can assume that a sum of the type

$$(4.7) \qquad \sum_{\vec{Q} \in \vec{\mathbf{Q}}} \widehat{\phi_{Q_1,1}}(\xi_1) \widehat{\phi_{Q_2,2}}(\xi_2) \widehat{\phi_{Q_3,3}}(\xi_3) \widehat{\phi_{Q_4,4}}(-\xi_1 - \xi_2 - \xi_3)$$

runs over a sparse collection of tiles $\vec{\mathbf{Q}}$. In such a sparse collection, then for every $Q \in \vec{\mathbf{Q}}$ there exists a unique shifted cube $\tilde{Q}$ in $\mathbb{R}^4$ such that $Q \subseteq \frac{7}{10}\tilde{Q}$ and with the diameter of Q similar to the diameter of $\tilde{Q}$. This allows us to assume that a sum of the type (4.7) runs over a sparse collections of shifted dyadic cubes such that $|Q_1| \sim |Q_2| \sim |Q_3| \sim |Q_4|$. Let $|\vec{Q}| \sim |Q_i|$, $i = 1, 2, 3, 4$, be the scale of the dyadic cube.

Further we know that in all three cases (4.2), (4.3) and (4.4) then the scale $|\vec{Q}|$ fixes the location of the tile Q_1. Also in the case (4.2) where we are close to the line $\beta\xi_2 + \xi_3 = 0$ then the tiles Q_2 and Q_3 can be made to overlap while in the second two cases (4.3), (4.4), when we are away from the line $\beta\xi_2 + \xi_3 = 0$ then Q_2 and Q_3 can be made to be a couple of units of length $|\vec{Q}|$ away from another so they don't overlap.

We will now study the quadlinear form associated to (3.4).

$$(4.8) \quad \int_{\mathbb{R}} \tilde{T}_\beta(f_1, f_2, f_3)(x) f_4(x) dx$$

$$= \int_{\xi_1 + \xi_2 + \xi_3 + \xi_4 = 0} \left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1 + \beta\xi_2 + \xi_3) d\alpha \right] \widehat{f_1}(\xi_1) \widehat{f_2}(\xi_2) \widehat{f_3}(\xi_3) \widehat{f_4}(\xi_4) d\xi_1 d\xi_2 d\xi_3 d\xi_4$$

$$= \sum_{\vec{Q} \in \vec{\mathbf{Q}}} \int_{\xi_1 + \xi_2 + \xi_3 + \xi_4 = 0} \frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1 + \beta\xi_2 + \xi_3) d\alpha \right]}{a(\xi_1, \xi_2, \xi_3)} \widehat{\phi_{Q_1,1}}(\xi_1) \widehat{\phi_{Q_2,2}}(\xi_2) \widehat{\phi_{Q_3,3}}(\xi_3) \widehat{\phi_{Q_4,4}}(-\xi_1 - \xi_2 - \xi_3)$$

$$\widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)\widehat{f_4}(\xi_4)d\xi_1 d\xi_2 d\xi_3 d\xi_4$$

$$= \sum_{\vec{Q}\in\vec{\mathbf{Q}}} \int_{\xi_1+\xi_2+\xi_3+\xi_4=0} \frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]}{a(\xi_1,\xi_2,\xi_3)}\widehat{\phi_{Q_1,1}}(\xi_1)\widehat{\phi_{Q_2,2}}(\xi_2)\widehat{\phi_{Q_3,3}}(\xi_3)\widehat{\phi_{Q_4,4}}(\xi_4)$$

$$\widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)\widehat{f_4}(\xi_4)d\xi_1 d\xi_2 d\xi_3 d\xi_4$$

We can write

$$\frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]}{a(\xi_1,\xi_2,\xi_3)}\widehat{\phi_{Q_1,1}}(\xi_1)\widehat{\phi_{Q_2,2}}(\xi_2)\widehat{\phi_{Q_3,3}}(\xi_3)$$

as

$$\frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]}{a(\xi_1,\xi_2,\xi_3)}\widehat{\tilde{\phi}_{Q_1,1}}(\xi_1)\widehat{\tilde{\phi}_{Q_2,2}}(\xi_2)\widehat{\tilde{\phi}_{Q_3,3}}(\xi_3)\widehat{\phi_{Q_1,1}}(\xi_1)\widehat{\phi_{Q_2,2}}(\xi_2)\widehat{\phi_{Q_3,3}}(\xi_3)$$

where $\widehat{\tilde{\phi}_{Q_1,1}}\otimes\widehat{\tilde{\phi}_{Q_2,2}}\otimes\widehat{\tilde{\phi}_{Q_3,3}}$ is identically equal to 1 on the support of $\widehat{\phi_{Q_1,1}}\otimes\widehat{\phi_{Q_2,2}}\otimes\widehat{\phi_{Q_3,3}}$.

Now split

$$\frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]}{a(\xi_1,\xi_2,\xi_3)}\widehat{\tilde{\phi}_{Q_1,1}}(\xi_1)\widehat{\tilde{\phi}_{Q_2,2}}(\xi_2)\widehat{\tilde{\phi}_{Q_3,3}}(\xi_3)$$

as a Fourier series

$$\sum_{n_1,n_2,n_3} C_{n_1,n_2,n_3}^{\vec{Q}} e^{2\pi i\frac{n_1}{|Q|}\xi_1}e^{2\pi i\frac{n_2}{|Q|}\xi_2}e^{2\pi i\frac{n_3}{|Q|}\xi_3}.$$

The coefficient $C_{n_1,n_2,n_3}^{\vec{Q}}$ is given by

$$(4.9)\quad C_{n_1,n_2,n_3}^{\vec{Q}} = \frac{1}{|\vec{Q}|^4}\int_{\mathbb{R}^3}\frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]}{a(\xi_1,\xi_2,\xi_3)}\widehat{\tilde{\phi}_{Q_1,1}}(\xi_1)\widehat{\tilde{\phi}_{Q_2,2}}(\xi_2)\widehat{\tilde{\phi}_{Q_3,3}}(\xi_3)$$
$$e^{-2\pi i\frac{n_1}{|Q|}\xi_1}e^{-2\pi i\frac{n_2}{|Q|}\xi_2}e^{-2\pi i\frac{n_3}{|Q|}\xi_3}d\xi_1 d\xi_2 d\xi_3$$

Lemma 4.7.
$$|C_{n_1,n_2,n_3}^{\vec{Q}}| \lesssim C(n_1,n_2,n_3)$$
where the implicit constant does not depend on $\vec{Q}$.

This lemma is a consequence of lemma 6.2 that we prove in section 6. The main point for now is that the Fourier coefficient is bounded uniformly independently of the dyadic cube $\vec{Q}$.

We can now majorize (4.8) by

$$\sum_{n_1,n_2,n_3} C(n_1,n_2,n_3)\sum_{\vec{Q}\in\vec{\mathbf{Q}}}|\int_{\xi_1+\xi_2+\xi_3+\xi_4=0}\widehat{\phi_{Q_1,1}}(\xi_1)\widehat{\phi_{Q_2,2}}(\xi_2)\widehat{\phi_{Q_3,3}}(\xi_3)\widehat{\phi_{Q_4,4}}(\xi_4)$$

$$\widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)\widehat{f_4}(\xi_4)e^{-2\pi i\frac{n_1}{|Q|}\xi_1}e^{-2\pi i\frac{n_2}{|Q|}\xi_2}e^{-2\pi i\frac{n_3}{|Q|}\xi_3}d\xi_1 d\xi_2 d\xi_3 d\xi_4|$$

$$= \sum_{n_1,n_2,n_3} C(n_1,n_2,n_3) \sum_{\vec{Q} \in \vec{\mathbf{Q}}} |\int_{\mathbb{R}^5} \widehat{f_1 * \phi_{Q_1,1}}(\xi_1) \widehat{f_2 * \phi_{Q_2,2}}(\xi_2) \widehat{f_3 * \phi_{Q_3,3}}(\xi_3) \widehat{f_4 * \phi_{Q_4,4}}(\xi_4)$$

$$e^{2\pi i(\xi_1+\xi_2+\xi_3+\xi_4)x} d\xi_1 d\xi_2 d\xi_3 d\xi_4 dx|$$

$$= \sum_{n_1,n_2,n_3} C(n_1,n_2,n_3) \sum_{\vec{Q} \in \vec{\mathbf{Q}}} |\int_{\mathbb{R}} (f_1 * \phi_{Q_1,1}^{n_1})(x)(f_2 * \phi_{Q_2,2}^{n_2})(x)(f_3 * \phi_{Q_3,3}^{n_3})(x)(f_4 * \phi_{Q_4,4})(x)dx|$$

Here the meaning of $\phi_{Q_i,i}^{n_i}$ is that if $\phi_{Q_i,i}$ was an L^1 normalized wave packet on $I_{\vec{Q}} \times Q_i$ then $\phi_{Q_i,i}^{n_i}$ is an L^1 normalized wave packet on $I_{\vec{Q}}^{n_i} \times Q_i$ where $I_{\vec{Q}}^{n_i}$ is a dyadic interval sitting n_i units of length $|I_{\vec{Q}}|$ away from $I_{\vec{Q}}$.

Split $\vec{\mathbf{Q}} = \bigcup_{k \in \mathbb{Z}} \vec{\mathbf{Q}}_k$ where $\vec{\mathbf{Q}}_k$ has cubes $\vec{Q}$ of scale $|\vec{Q}| = 2^k$ and thus $|I_{\vec{Q}}| = 2^{-k}$.

$$(4.10) \quad \sum_{\vec{Q} \in \vec{\mathbf{Q}}} |\int_{\mathbb{R}} (f_1 * \phi_{Q_1,1}^{n_1})(x)(f_2 * \phi_{Q_2,2}^{n_2})(x)(f_3 * \phi_{Q_3,3}^{n_3})(x)(f_4 * \phi_{Q_4,4})(x)dx|$$

$$= \sum_{k \in \mathbb{Z}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_k} |2^{-k} \int_{\mathbb{R}} (f_1 * \phi_{Q_1,1}^{n_1})(2^{-k}y)(f_2 * \phi_{Q_2,2}^{n_2})(2^{-k}y)(f_3 * \phi_{Q_3,3}^{n_3})(2^{-k}y)(f_4 * \phi_{Q_4,4})(2^{-k}y)dy|$$

$$= \sum_{k \in \mathbb{Z}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_k} \left| |I_{\vec{Q}}| \int_0^1 \sum_{m \in \mathbb{Z}} (f_1 * \phi_{Q_1,1}^{n_1})(2^{-k}m + 2^{-k}\gamma)(f_2 * \phi_{Q_2,2}^{n_2})(2^{-k}m + 2^{-k}\gamma) \right.$$

$$\left. (f_3 * \phi_{Q_3,3}^{n_3})(2^{-k}m + 2^{-k}\gamma)(f_4 * \phi_{Q_4,4})(2^{-k}m + 2^{-k}\gamma)d\gamma \right|$$

Now observe that for $i = 1,2,3,4$ (where we take $n_4 = 0$)

$$(f_i * \phi_{Q_i,i}^{n_i})(2^{-k}m + 2^{-k}\gamma) = \int_{\mathbb{R}} f_i(z)\phi_{Q_i,i}^{n_i}(2^{-k}m + 2^{-k}\gamma - z)dz$$

$$= \frac{1}{|I_{\vec{Q}}|^{1/2}} \int_{\mathbb{R}} f_i(z)|I_{\vec{Q}}|^{1/2}\phi_{Q_i,i}^{n_i}(2^{-k}m + 2^{-k}\gamma - z)dz$$

$$= \frac{1}{|I_{\vec{Q}}|^{1/2}} \langle f_i, \tilde{\phi}_{Q_i,i,m,\gamma}^{n_i} \rangle$$

where $\tilde{\phi}_{Q_i,i,m,\gamma}^{n_i}$ is a wave packet translated from $\phi_{Q_i,i}^{n_i}$ by m steps in time and then additionally shifted by γ steps. Note that $\tilde{\phi}_{Q_i,i,m,\gamma}^{n_i}$ is an L^2 normalized wave packet since $\phi_{Q_i,i}^{n_i}$ was L^1 normalized. Now (4.10) becomes

$$\int_0^1 \sum_{\vec{Q} \in \vec{\mathbf{Q}}} \sum_{m \in \mathbb{Z}} |I_{\vec{Q}}| \frac{1}{|I_{\vec{Q}}|^{1/2}} \langle f_1, \tilde{\phi}_{Q_1,1,m,\gamma}^{n_1} \rangle \frac{1}{|I_{\vec{Q}}|^{1/2}} \langle f_2, \tilde{\phi}_{Q_2,2,m,\gamma}^{n_2} \rangle \frac{1}{|I_{\vec{Q}}|^{1/2}} \langle f_3, \tilde{\phi}_{Q_3,3,m,\gamma}^{n_3} \rangle$$

$$\frac{1}{|I_{\vec{Q}}|^{1/2}} \langle f_4, \tilde{\phi}_{Q_4,4,m,\gamma} \rangle d\gamma$$

$$= \int_0^1 \sum_{\vec{P} \in \vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|} \langle f_1, \tilde{\phi}_{P_1^{n_1},1,\gamma} \rangle \langle f_2, \tilde{\phi}_{P_2^{n_2},2,\gamma} \rangle \langle f_3, \tilde{\phi}_{P_3^{n_3},3,\gamma} \rangle \langle f_4, \tilde{\phi}_{P_4,4,\gamma} \rangle d\gamma$$

where $P_i^{n_i}$ denotes the tile $I_{P_i}^{n_i+m} \times Q_i$ where $I_{P_i}^{n_i+m}$ is a dyadic interval such that $|I_{P_i}^{n_i+m}| \sim |Q_i|^{-1}$ for $i = 1, 2, 3, 4$ (again we have $n_4 = 0$). Again then $I_{P_i}^{n_i+m}$ sits $n_i + m$ units of length $|I_{\vec{P}}|$ away from I_{P_i}.

If we now fix $n_1, n_2, n_3 \in \mathbb{Z}$ and $\gamma \in [0, 1]$ then it is sufficient to study estimates for the following discrete variant of (4.8)

$$\sum_{\vec{P} \in \mathbf{P}} \frac{1}{|I_{\vec{P}}|} \langle f_1, \phi_{P_1^{n_1},1} \rangle \langle f_2, \phi_{P_2^{n_2},2} \rangle \langle f_3, \phi_{P_3^{n_3},3} \rangle \langle f_4, \phi_{P_4,4} \rangle$$

Write

$$(4.11) \qquad \Lambda_{\mathbf{P}}(f_1, f_2, f_3, f_4) := \sum_{\vec{P} \in \mathbf{P}} \frac{1}{|I_{\vec{P}}|} \langle f_1, \phi_{P_1^{n_1},1} \rangle \langle f_2, \phi_{P_2^{n_2},2} \rangle \langle f_3, \phi_{P_3^{n_3},3} \rangle \langle f_4, \phi_{P_4,4} \rangle$$

and define $T_{\vec{\mathbf{P}}}(f_1, f_2, f_3)$ with

$$\langle T_{\vec{\mathbf{P}}}(f_1, f_2, f_3), f_4 \rangle = \Lambda_{\vec{\mathbf{P}}}(f_1, f_2, f_3, f_4)$$

To compare our quadtiles with the tiles one faces in the bilinear Hilbert transform then notice that if $\vec{P} = (P_1, P_2, P_3, P_4)$ then P_1 is like a paraproduct tile, P_2 and P_3 might at a first glance seem just as in the bilinear Hilbert transform and P_4 is essentially as in the bilinear Hilbert transform, just potentially translated a bit in frequency by P_1. Note that the constant in the definition of $\lesssim'$ is 5 as opposed to 3 in [13]. We choose a bigger constant to make up for this extra possible translation of P_4. In the next section we will see in which cases we are essentially as in the bilinear Hilbert transform case, and in which cases we have to be more careful.

5. RANK $(1, 0)$

Recall a standard definition of rank [13].

Definition 5.1. *A collection $\vec{\mathbf{P}}$ of quadtiles is said to have rank 1 if one has the following properties for all $\vec{P}, \vec{P}' \in \vec{\mathbf{P}}$:*

- *If $\vec{P} \neq \vec{P}'$, then $P_j \neq P'_j$ for all j=1,2,3,4.*
- *If $P'_j \leq P_j$ for some $j = 1, 2, 3, 4$, then $P'_i \lesssim P_i$ for all $1 \leq i \leq 4$.*
- *If we further assume that $10^9 |I_{\vec{P}'}| < |I_{\vec{P}}|$, then we have $P'_i \lesssim' P_i$ for all $i \neq j$.*

This definition does not work for our collection of quadtiles because the paraproduct tile P_1 does not uniquely determine the other three tiles.

We only need a frequency or time interval from one of our tiles to determine P_1, while we need a whole tile P_j, $j = 2, 3$ or 4, to determine the other three. Motivated by this fact and what ingredients are really important in a rank definition [15] we give the following definition.

Definition 5.2. *Let $\{i_1, i_2, i_3, i_4\}$ be some rearrangement of $\{1, 2, 3, 4\}$. A collection $\vec{\mathbf{P}}$ of quadtiles is said to have rank $(1, 0)$ with respect to $\{\{i_1, i_2, i_3\}, \{i_4\}\}$ if one has the following properties for all $\vec{P}, \vec{P}' \in \vec{\mathbf{P}}$:*

- *If $\vec{P} \neq \vec{P}'$, then $P_{i_j} \neq P'_{i_j}$ for all j=1,2,3 and if $I_{\vec{P}} = I_{\vec{P}'}$, then $P_{i_4} = P'_{i_4}$.*
- *If $P'_{i_j} \leq P_{i_j}$ for some $j = 1, 2, 3$, then $P'_{i_k} \lesssim P_{i_k}$ for all $1 \leq k \leq 4$.*
- *If we further assume that $10^9 |I_{\vec{P}'}| < |I_{\vec{P}}|$, then there exist at least two indices*

$$\tau_1(i_j), \tau_2(i_j) \in \{1, 2, 3, 4\} \setminus \{i_j\}, \quad \tau_1(i_j) \neq \tau_2(i_j)$$

such that we have $P'_{\tau_1(i_j)} \lesssim' P_{\tau_1(i_j)}$ and $P'_{\tau_2(i_j)} \lesssim' P_{\tau_2(i_j)}$. We call those indices good indices with respect to i_j and note that there might be up to three of them. Here we understand $P'_{i_4} \lesssim' P_{i_4}$ to mean $\omega_{P'_{i_4}} \cap \omega_{P_{i_4}} = \emptyset$.

Note that the orderings $\leq$ and $\lesssim'$ do not make sense for our paraproduct tiles because we have the relation $\leq$ between any two such tiles and thus $\lesssim'$ never happens. These orderings work well on the bilinear Hilbert transform type tiles where flexibility is helpful. We have to be more exact with the paraproduct tiles and thus understand the relation $\leq$ to mean that the paraproduct tiles intersect in frequency while $\lesssim'$ means that they don't intersect.

It is not hard to see that our collection of quadtiles is rank $(1,0)$ with respect to $\{\{2,3,4\},\{1\}\}$ where a collection corresponds to exactly one of the three cases we have. The first and second conditions are clearly fulfilled since knowing one of the bilinear Hilbert transform tiles gives us complete information about all the other tiles and since the paraproduct tile is completely determined by the time interval. Modulo a finite refinement of our collection we can also see that the last condition is fulfilled.

Assume we are in the case (4.2) and that we have $10^9|I_{\vec{P}'}| < |I_{\vec{P}}|$ and $P'_2 \leq P_2$. We cannot guarantee that $P'_3 \leq P_3$ since P_2 and P_3 are essentially the same tile and similarly for P'_2 and P'_3. However $10^9|I_{\vec{P}'}| < |I_{\vec{P}}|$ guarantees that $\omega_{P'_1} \cap \omega_{P_1} = \emptyset$ which along with the previous observation also guarantees that $P'_4 \lesssim' P_4$. The other possibilities in this case go somewhat similarly. This particular example shows how critical the paraproduct tile is in our analysis.

In the case (4.3) then P_1 has minimal effect so we are essentially in the bilinear Hilbert case so all the conditions above are fulfilled.

Assume we are in the case (4.4) and that we have $10^9|I_{\vec{P}'}| < |I_{\vec{P}}|$ and $P'_4 \leq P_4$. We claim that $P'_2 \lesssim' P_2$ and $P'_3 \lesssim' P_3$ so let us assume for contradiction that $P'_2 \leq P_2$. The distance between the centers of the frequency supports of P_1 and P'_1 is roughly $|\omega_{P'_1}| - |\omega_{P_1}| < |\omega_{P'_1}|$ which means, since $P'_2 \leq P_2$ and $P'_4 \leq P_4$, that the distance between the centers of the frequency supports of P_3 and P'_3 is at most $|\omega_{P'_1}|$ which gives $P'_3 \leq P_3$. This must be a contradiction and thus we have $P'_2 \lesssim' P_2$ and $P'_3 \lesssim' P_3$. The other possibilities in this case go somewhat similarly.

6. FOURIER COEFFICIENT

Recall from (4.9) that the Fourier coefficient $C^{\vec{Q}}_{n_1,n_2,n_3}$ is given by

$$C^{\vec{Q}}_{n_1,n_2,n_3} = \frac{1}{|\vec{Q}|^4} \int_{\mathbb{R}^3} \frac{\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]}{a(\xi_1,\xi_2,\xi_3)} \widehat{\phi_{Q_1,1}}(\xi_1)\widehat{\phi_{Q_2,2}}(\xi_2)\widehat{\phi_{Q_3,3}}(\xi_3)$$
$$e^{-2\pi i\frac{n_1}{|\vec{Q}|}\xi_1}e^{-2\pi i\frac{n_2}{|\vec{Q}|}\xi_2}e^{-2\pi i\frac{n_3}{|\vec{Q}|}\xi_3}d\xi_1 d\xi_2 d\xi_3$$

Change variables and obtain

$$C^{\vec{Q}}_{n_1,n_2,n_3} = \int_{\mathbb{R}^3}\left[\int_0^1 1_{\mathbb{R}_+}(\alpha\xi_1+\beta\xi_2+\xi_3)d\alpha\right]\frac{\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(\xi_3)}{\tilde{a}(\xi_1,\xi_2,\xi_3)}e^{-2\pi in_1\xi_1}$$
$$e^{-2\pi in_2\xi_2}e^{-2\pi in_3\xi_3}d\xi_1 d\xi_2 d\xi_3$$

where $\widehat{\phi_i}(\xi_i) = \widehat{\phi_{Q_1,1}}(|\vec{Q}|\xi_1)$ is a bump that is of scale 1 and $\tilde{a}(\xi_1,\xi_2,\xi_3) = a(|\vec{Q}|\xi_1,|\vec{Q}|\xi_2,|\vec{Q}|\xi_3)$ is also of scale 1 on the support of $\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(\xi_3)$. To see why the last statement is true we have to recall

$$a(\xi_1, \xi_2, \xi_3) = \sum_{\vec{\tilde{Q}} \in \vec{\mathbf{Q}}} \widehat{\phi_{\tilde{Q}_1, 1}}(\xi_1) \widehat{\phi_{\tilde{Q}_2, 2}}(\xi_2) \widehat{\phi_{\tilde{Q}_3, 3}}(\xi_3)$$

and split into cases based on (4.2), (4.3) and (4.4). First note that for a term in

$$\sum_{\vec{\tilde{Q}} \in \vec{\mathbf{Q}}} \widehat{\phi_{\tilde{Q}_1, 1}}(\xi_1) \widehat{\phi_{\tilde{Q}_2, 2}}(\xi_2) \widehat{\phi_{\tilde{Q}_3, 3}}(\xi_3)$$

to contribute to the sum on the support of $\vec{Q}$ we must have $\tilde{Q}_i \cap Q_i \neq \emptyset$ for $i = 1, 2, 3$.

Start with the cases (4.2) and (4.4). For $\tilde{Q}_1 \cap Q_1 \neq \emptyset$ we must have $|\vec{\tilde{Q}}| \sim |\vec{Q}|$ because else $\widehat{\phi_{\tilde{Q}_1, 1}}$ and $\widehat{\phi_{Q_1, 1}}$ have disjoint supports.

The last case is (4.3). Assume we have $\vec{\tilde{Q}}$ and $\vec{Q}$ such that $\tilde{Q}_i \cap Q_i \neq \emptyset$ for $i = 1, 2, 3$. Let's now for symmetry assume we have $|\vec{\tilde{Q}}| \ll |\vec{Q}|$. We are in the case where Q_2 and Q_3 are several units of length $|\vec{Q}|$ away from one another and $\tilde{Q}_2$ and $\tilde{Q}_3$ are several units of length $|\vec{\tilde{Q}}|$ away from one another. However if $\tilde{Q}_2 \cap Q_2 \neq \emptyset$ then we can't have $\tilde{Q}_3 \cap Q_3 \neq \emptyset$ which is a contradiction. Thus we must have $|\vec{\tilde{Q}}| \sim |\vec{Q}|$.

We now want to integrate by parts to obtain decay in n_1, n_2, n_3. We do not need to worry about derivatives hitting $\frac{\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(\xi_3)}{\tilde{a}(\xi_1, \xi_2, \xi_3)}$ which is smooth and of scale 1.

In the case (4.3) we do not catch the planes where our symbol is continuous but not differentiable. In that case we can thus integrate by parts as often as we want and obtain as much decay in n_1, n_2 and n_3 as we want.

In the other cases, (4.2) and (4.3), we might catch the planes where our symbol is merely continuous but in both cases we know that Q_1 is away from the origin. Thus we can write $C_{n_1, n_2, n_3}^{\vec{Q}}$ as

$$(6.1) \quad \int_{\mathbb{R}^3} \left[\frac{1}{\xi_1} \int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right] \frac{\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(\xi_3)}{\tilde{a}(\xi_1, \xi_2, \xi_3)} e^{-2\pi i n_1 \xi_1}$$
$$e^{-2\pi i n_2 \xi_2} e^{-2\pi i n_3 \xi_3} d\xi_1 d\xi_2 d\xi_3$$

$$= \int_{\mathbb{R}^3} \left[\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right] \frac{\widehat{\tilde{\phi}_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(\xi_3)}{\tilde{a}(\xi_1, \xi_2, \xi_3)} e^{-2\pi i n_1 \xi_1}$$
$$e^{-2\pi i n_2 \xi_2} e^{-2\pi i n_3 \xi_3} d\xi_1 d\xi_2 d\xi_3$$

where $\widehat{\tilde{\phi}_1}(\xi_1) = \frac{1}{\xi_1}\widehat{\phi_1}(\xi_1)$ is well defined and still smooth because ξ_1 is always away from zero. As in Muscalu's treatment of the symbol for the Calderón commutator [11], which has a non-standard symbol, we get the following lemmas.

Lemma 6.1. *One has the following identities*

a) $\partial_{\xi_3}^2 \left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \delta_0(\xi_1 + \beta\xi_2 + \xi_3) - \delta_0(\beta\xi_2 + \xi_3)$

b) $\partial_{\xi_2} \partial_{\xi_3} \left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \beta(\delta_0(\xi_1 + \beta\xi_2 + \xi_3) - \delta_0(\beta\xi_2 + \xi_3))$

c) $\partial_{\xi_1} \partial_{\xi_3} \left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \delta_0(\xi_1 + \beta\xi_2 + \xi_3)$

d) $\partial_{\xi_2}^2 \left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \beta^2(\delta_0(\xi_1 + \beta\xi_2 + \xi_3) - \delta_0(\beta\xi_2 + \xi_3))$

e) $\partial_{\xi_1} \partial_{\xi_2} \left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \beta \delta_0(\xi_1 + \beta\xi_2 + \xi_3)$

f) $\partial_{\xi_1}^2 \left(\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \delta_0(\xi_1 + \beta\xi_2 + \xi_3)$

Proof. This is straightforward. Let us verify a) for instance. One has

$$
\partial_{\xi_3}^2 \left(\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha \right) = \partial_{\xi_3} \left(\int_0^{\xi_1} \delta_0(\alpha + \beta\xi_2 + \xi_3) d\alpha \right)
$$
$$
= \partial_{\xi_3} \left(\int_{\beta\xi_2 + \xi_3}^{\xi_1 + \beta\xi_2 + \xi_3} \delta_0(\alpha) d\alpha \right)
$$
$$
= \delta_0(\xi_1 + \beta\xi_2 + \xi_3) - \delta_0(\beta\xi_2 + \xi_3)
$$

$\square$

Lemma 6.2.

$$
|C_{n_1,n_2,n_3}^{\vec{Q}}| \lesssim c_\beta^1 \frac{1}{\langle n_3 \rangle^2} \cdot \frac{1}{\langle n_1 - n_3 \rangle^{M_2}} \cdot \frac{1}{\langle n_2 - \beta n_3 \rangle^{M_3}} + c_\beta^2 \frac{1}{\langle n_3 \rangle^2} \cdot \frac{1}{\langle n_1 \rangle^{M_2}} \cdot \frac{1}{\langle n_2 - \beta n_3 \rangle^{M_3}}
$$
$$
+ c_\beta^3 \frac{1}{\langle n_3 \rangle^{M_1}} \cdot \frac{1}{\langle n_1 - \frac{n_2}{\beta} \rangle^{M_2}} \cdot \frac{1}{\langle n_3 - \frac{n_2}{\beta} \rangle^{M_3}} + c_\beta^4 \frac{1}{\langle n_3 \rangle^{M_1}} \cdot \frac{1}{\langle n_1 \rangle^{M_2}} \cdot \frac{1}{\langle n_3 - \frac{n_2}{\beta} \rangle^{M_3}}
$$
$$
+ c_\beta^5 \frac{1}{\langle n_3 \rangle^{M_1}} \cdot \frac{1}{\langle n_2 \rangle^{M_2}} \cdot \frac{1}{\langle n_3 - n_1 \rangle^{M_3}} \cdot \frac{1}{\langle n_2 - \beta n_1 \rangle^{M_4}} + c_\beta^6 \frac{1}{\langle n_3 \rangle^{M_1}} \cdot \frac{1}{\langle n_2 \rangle^{M_2}} \cdot \frac{1}{\langle n_1 \rangle^{M_3}}
$$

where $\langle n \rangle := 2 + |n|$ and M_1, M_2, M_3, M_4 are fixed large integers and $c_\beta^1, \ldots, c_\beta^6$ are constants that only depend on β.

Proof. As mentioned before then this clearly holds in the case (4.3) since then the symbol is smooth and we can integrate by parts as often as we want in the Fourier coefficient. In the other two cases (4.2) and (4.4) we must use lemma 6.1. The idea is to integrate by parts in (6.1) in the ξ_3 variable as often as we can. Since both $\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha$ and $\frac{\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(\xi_3)}{\tilde{a}(\xi_1,\xi_2,\xi_3)}$ depend on ξ_3 then derivatives can hit either of the terms. If the derivative hits the term $\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha$ twice then because of lemma 6.1 the ξ_3 variable disappears and (6.1) collapses to

$$
\int_{\mathbb{R}^2} \frac{\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(-\xi_1 - \beta\xi_2)}{\tilde{a}(\xi_1, \xi_2, -\xi_1 - \beta\xi_2)} e^{-2\pi i(n_1 - n_3)\xi_1} e^{-2\pi i(n_2 - \beta n_3)\xi_2} d\xi_1 d\xi_2
$$
$$
- \int_{\mathbb{R}^2} \frac{\widehat{\phi_1}(\xi_1)\widehat{\phi_2}(\xi_2)\widehat{\phi_3}(-\beta\xi_2)}{\tilde{a}(\xi_1, \xi_2, -\xi_1 - \beta\xi_2)} e^{-2\pi i n_1 \xi_1} e^{-2\pi i(n_2 - \beta n_3)\xi_2} d\xi_1 d\xi_2
$$

The integrands in both those terms are smooth and can be integrated by parts as many times as we wish and all the derivatives are compactly supported on scale 1. This explains the appearance of the first two terms in the estimate for $C_{n_1,n_2,n_3}^{\vec{Q}}$.

If however the ξ_3 derivative didn't hit the term $\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha$ two times, even after running the procedure many times, this means that we already gained a factor of the type $\frac{1}{\langle n_3 \rangle^{M_1}}$, at which point we stop integrating by parts in ξ_3 and start integrating by parts in ξ_2. If ξ_2 derivatives hit $\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha$ we face two possible cases, we either end up with $\partial_{\xi_2}\partial_{\xi_3} \left(\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha \right)$ or $\partial_{\xi_2}^2 \left(\int_0^{\xi_1} 1_{\mathbb{R}_+} (\alpha + \beta\xi_2 + \xi_3) d\alpha \right)$. Using lemma 6.1 then the integral collapses as in the first case, that is ξ_2 becomes $-\frac{\xi_3 + \xi_1}{\beta}$ or $-\frac{\xi_3}{\beta}$. After that

we are, as before, integrating by parts a smooth function, obtaining an upper bound that explains the appearance of the third and fourth terms in the estimate for $C_{n_1,n_2,n_3}^{\vec{Q}}$.

If however $\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha+\beta\xi_2+\xi_3)d\alpha$ has not been hit two times by some combination of ξ_3 and ξ_2 derivatives after running the procedure many times, this means that we have already gained a factor of the type $\frac{1}{\langle n_3 \rangle^{M_1}} \cdot \frac{1}{\langle n_2 \rangle^{M_2}}$ at which point we stop integrating by parts in ξ_2 and start integrating by parts in ξ_1. If ξ_1 derivatives hit $\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha+\beta\xi_2+\xi_3)d\alpha$ we face three possible cases, we end up with $\partial_{\xi_1}\partial_{\xi_3}\left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha+\beta\xi_2+\xi_3)d\alpha\right)$, $\partial_{\xi_1}\partial_{\xi_2}\left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha+\beta\xi_2+\xi_3)d\alpha\right)$ or $\partial_{\xi_1}^2\left(\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha+\beta\xi_2+\xi_3)d\alpha\right)$. Using lemma 6.1 the integral collapses as before, that is ξ_1 becomes $-\beta\xi_2-\xi_3$. After that we are, as before, integrating by parts a smooth function, obtaining an upper bound that explains the appearance of the fifth term in the estimate for $C_{n_1,n_2,n_3}^{\vec{Q}}$.

Last but not least, if no combination of ξ_1, ξ_2 or ξ_3 derivatives hits $\int_0^{\xi_1} 1_{\mathbb{R}_+}(\alpha+\beta\xi_2+\xi_3)d\alpha$ twice then this means that the derivatives keep hitting the smooth function in which case we obtain an upper bound that explains the appearance of the last term in the estimate for $C_{n_1,n_2,n_3}^{\vec{Q}}$.

$\square$

7. Discrete Operator

Let now $\vec{\mathbf{P}}$ be a finite collection of multitiles which is sparse and has rank $(1,0)$. Consider also wave packets $(\phi_{P_j^{n_j},j})_{\vec{P}\in\vec{\mathbf{P}}}$ for $j=1,2,3,4$ adapted to the tiles $P_j^{n_j}$ respectively as before where n_1, n_2 and n_3 are fixed and $n_4=0$. Assume also that they are all L^2-normalized. The following theorem will be proven in detail in section 9.

Theorem 7.1. *Let γ_1 and γ_3 be positive numbers, smaller than 1 but very close to 1, γ_2 be a positive number smaller than $\frac{1}{2}$ but very close to $\frac{1}{2}$. Let also E_1, E_2, E_3, $E_4 \subseteq \mathbb{R}$ be measurable sets of finite measure. Then there exists $E_4' \subseteq E_4$ with $|E_4'| \sim |E_4|$ such that for every $|f_1| \lesssim 1_{E_1}$, $|f_2| \lesssim 1_{E_2}$, $|f_3| \lesssim 1_{E_3}$ one has*

$$(7.1)\qquad \left|\int_{\mathbb{R}} T_{\vec{\mathbf{P}}}(f_1,f_2,f_3)(x)1_{E_4'}(x)dx\right| \lesssim \left(\prod_{j=1}^3 |\log_2(\langle n_j\rangle)|^4\right) |E_1|^{\gamma_1}|E_2|^{\gamma_2}|E_3|^{\gamma_3}|E_4|^{\gamma_4}$$

where γ_4 is defined by $\gamma_1+\gamma_2+\gamma_3+\gamma_4=1$. Moreover the implicit constant is independent of the cardinality of $\vec{\mathbf{P}}$.

Using the interpolation theory by Muscalu, Tao and Thiele [15], the symmetries of $T_{\vec{\mathbf{P}}}$ and standard duality arguments then one can deduce the following theorem.

Theorem 7.2. *If $\vec{\mathbf{P}}$ is as before then $T_{\vec{\mathbf{P}}}$ maps boundedly*

$$(7.2)\qquad\qquad T_{\vec{\mathbf{P}}} : L^{p_1}(\mathbb{R}) \times L^{p_2}(\mathbb{R}) \times L^{p_3}(\mathbb{R}) \mapsto L^{p_4}(\mathbb{R})$$

for any $1 < p_1,p_2,p_3 \leq \infty$ and $\frac{2}{5} < p_4 < \infty$ such that $\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}=\frac{1}{p_4}$ and $\frac{2}{3} < \frac{p_2 p_3}{p_2+p_3} \leq \infty$. Furthermore, the constant of boundedness depends on n_1, n_2, n_3 and n_4 in a way that can be bounded by $\prod_{j=1}^3 |\log_2(\langle n_j\rangle)|^4$.

Note that this is a stronger result than in theorem 1.1.

To prove Theorem 1.1 then let p_1, p_2, p_3 and p_4 be as in the theorem and recall that in section 3 we commented that it is enough to show the theorem for $\tilde{T}_\beta$. If $p_4 \geq 1$ then standard arguments extend the theorem to $\tilde{T}_\beta$. If however $p_4 < 1$ let $f_i \in L^{p_i}(\mathbb{R})$, $i = 1, 2, 3$ and note

$$
\begin{aligned}
\|\tilde{T}_\beta(f_1, f_2, f_3)\|_{p_4} &= \|\tilde{T}_\beta(f_1, f_2, f_3)^{p_4}\|_1^{1/p_4} \\
&\lesssim \| \left(\int_0^1 \sum_{n_1, n_2, n_3} C(n_1, n_2, n_3) T_{\vec{\mathbf{P}}, \eta}(f_1, f_2, f_3) d\eta \right)^{p_4} \|_1^{1/p_4} \\
&\lesssim \| \int_0^1 \sum_{n_1, n_2, n_3} C(n_1, n_2, n_3)^{p_4} T_{\vec{\mathbf{P}}, \eta}^{p_4}(f_1, f_2, f_3) d\eta \|_1^{1/p_4}
\end{aligned}
$$

This last step is only well defined if $p_4 > \frac{1}{2}$ because $C(n_1, n_2, n_3)$ includes terms that contain $\frac{1}{\langle n_3 \rangle^2}$ by lemma 6.2 and we need $\frac{1}{\langle n_3 \rangle^{2p_4}}$ to be summable. In that case then theorem 7.2 and lemma 6.2, along with standard results on the convergence of series of the type $\sum_n \frac{|\log_2(\langle n \rangle)|^4}{\langle n \rangle^p}$ where $p > 1$, can be used to conclude that theorem 1.1 holds true for $\tilde{T}_\beta$ and thus for T_β.

Note that the reason why $p_4 > \frac{1}{2}$ might seem a bit naive. One could hope to improve the result by treating all the $T_{\vec{\mathbf{P}}}$ simultaneously by picking in section 10 a common exceptional set. Such a strategy leads to a loss of $\langle n \rangle^{(1+\epsilon)}$, $\epsilon > 0$, in the size estimates in section 9. Thus, running through the standard argument, one would eventually have to control a sum of the following type

$$
\sum_{n_1, n_2, n_3} C(n_1, n_2, n_3) \langle n_1 \rangle^{(1+\epsilon)b_1} \langle n_2 \rangle^{(1+\epsilon)b_2} \langle n_3 \rangle^{(1+\epsilon)b_3}.
$$

If we only consider the first term in the estimate of the Fourier coefficient one faces the following sum

$$
\sum_{n_1, n_2, n_3} \frac{\langle n_3 \rangle^{(1+\epsilon)b_3}}{\langle n_3 \rangle^2} \cdot \frac{\langle n_1 \rangle^{(1+\epsilon)b_1}}{\langle n_1 - n_3 \rangle^{M_2}} \cdot \frac{\langle n_2 \rangle^{(1+\epsilon)b_2}}{\langle n_2 - \beta n_3 \rangle^{M_3}}.
$$

Changing variables through

$$
k = n_1 - n_3 \text{ and } l = n_2 - \beta n_3
$$

the sum becomes

$$
\sum_{k, l, n_3} \frac{\langle n_3 \rangle^{(1+\epsilon)b_3}}{\langle n_3 \rangle^2} \cdot \frac{\langle k + n_3 \rangle^{(1+\epsilon)b_1}}{\langle k \rangle^{M_2}} \cdot \frac{\langle l + \beta n_3 \rangle^{(1+\epsilon)b_2}}{\langle l \rangle^{M_3}}.
$$

Hence, one would like the expression

$$
\frac{\langle n_3 \rangle^{(1+\epsilon)(b_1 + b_2 + b_3)}}{\langle n_3 \rangle^2}.
$$

to be summable, which places stringent requirements on b_1, b_2 and b_3. In fact, if one goes thoroughly through the standard argument it is not hard to see that the condition $p_4 > \frac{1}{2}$ cannot be improved. It is thus an interesting open question whether this condition can be improved, which clearly either requires some novel ideas or some more delicate estimates.

8. TREES

The standard approach to prove the desired estimates for the form $\Lambda_{\vec{P}}$ is to organize the collection of quadtiles $\vec{P}$ into trees. We may assume, and will do so for the rest of the article, that $\vec{P}$ is sparse and of rank $(1,0)$. We will now recall basic definitions and comments for trees from [13]. The only change is that we will not consider 1 trees at all. We will essentially ignore the first position when setting up the trees. Also note that we set up the trees based on untranslated tiles.

Definition 8.1. *For any $2 \leq j \leq 4$ and a quadtile $\vec{P}_T \in \vec{P}$, define a j-tree with top $\vec{P}_T$ to be a collection of quadtiles $T \subseteq \vec{P}$ such that*

$$(8.1) \qquad\qquad\qquad P_j \leq P_{T,j} \text{ for all } \vec{P} \in T,$$

where $P_{T,j}$ is the j component of $\vec{P}_T$. We write I_T and $\omega_{T,j}$ for $I_{\vec{P}_T}$ and $\omega_{P_{T,j}}$ respectively. We say that T is a tree if it is a j-tree for some $2 \leq j \leq 4$.

Note that T does not necessarily have to contain its top $\vec{P}_T$.

Definition 8.2. *Let $2 \leq i \leq 4$. Two trees T, T' are said to be strongly i-disjoint if*

- *$P_i \neq P_i'$ for all $\vec{P} \in T$, $\vec{P}' \in T'$.*
- *Whenever $\vec{P} \in T$, $\vec{P}' \in T'$ are such that $2\omega_{P_i} \cap 2\omega_{P_i'} \neq \emptyset$, then one has $I_{\vec{P}'} \cap I_T = \emptyset$, and similarly with T and T' reversed.*

Note that if T and T' are strongly i-disjoint, then $I_P \times 2\omega_{P_i} \cap I_{P'} \times 2\omega_{P_i'} = \emptyset$ for all $\vec{P} \in T$, $\vec{P}' \in T'$.

Given that $\vec{P}$ is sparse, it is easy to see that if T is an i-tree, then for all $\vec{P}, \vec{P}' \in T$ and $j \neq i$, $2 \leq j \leq 4$, we have

$$\omega_{P_j} = \omega_{P_j'}$$

or

$$2\omega_{P_j} \cap 2\omega_{P_j'} = \emptyset$$

We pick trees for tiles $\vec{P}$ as in the bilinear Hilbert transform case but remember that our wave packets are in general adapted to tiles $P_i^{n_i}$, $i = 1, 2, 3$, that are translated in time by n_i units of length $|I_{\vec{P}}|$. Thus the effective trees we face are translated and are furthermore not evenly translated.

Due to the dyadic structure of the trees and the dyadic structure of the translation applied to the tiles in the trees then one can see that we can do better than saying that a translated tree, derived from a tree T, is supported on $\bigcup_{j=0}^{n_i} I_T^j$. As Muscalu observes [11] (and can be seen from the argument in section 11) then in fact the translated tree is supported on $\bigcup_{j \in Fr(n_i)} I_T^j$ where $Fr(n_i)$ is a set of indices that contains for example 0, 1 and n_i. We also know the following fact about the cardinality of $Fr(n_i)$

$$|Fr(n_i)| \lesssim \log_2(\langle n_i \rangle).$$

We call $\bigcup_{j \in Fr(n_i)} I_T^j$ "I_T and friends".

9. TILE NORMS

Let's recall the standard tile norms from the paper by Muscalu, Tao and Thiele [13].

Definition 9.1. *Let $\vec{\mathbf{P}}$ be a finite collection of quadtiles, $j = 1, 2, 3, 4$ and let $(a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}$ be a sequence of complex numbers. We define the size of this sequence by*

$$size_j((a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}) := \sup_{T \subset \vec{\mathbf{P}}} \left(\frac{1}{|I_T|} \sum_{\vec{P} \in T} |a_{P_j}|^2 \right)^{1/2}$$

where T ranges over all trees in $\vec{\mathbf{P}}$ which are either one quadtile trees or i-trees for some $2 \leq i \leq 4$ such that j is a good index with respect to i, as in the definition of rank $(1, 0)$.
We also define the energy of a sequence by

$$energy_j((a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}) := \sup_{n \in \mathbb{Z}} \sup_{\mathbb{T}} 2^n \left(\sum_{T \in \mathbf{T}} |I_T| \right)^{1/2}$$

where $\mathbf{T}$ ranges over all collections of strongly j-disjoint trees, $2 \leq j \leq 4$, in $\vec{\mathbf{P}}$ such that

$$\left(\sum_{\vec{P} \in T} |a_{P_j}|^2 \right)^{1/2} \geq 2^n |I_T|^{1/2}$$

for all $T \in \mathbf{T}$ and

$$\left(\sum_{\vec{P} \in T'} |a_{P_j}|^2 \right)^{1/2} \leq 2^{n+1} |I_{T'}|^{1/2}$$

for all sub-trees $T' \subset T \in \mathbf{T}$.

We will use those definitions for $a_{P_j} = \langle f_j, \phi_{P_j^{n_j}, j} \rangle$. Note that the restriction to i-trees for some $2 \leq i \leq 4$ such that j is a good index with respect to i, as in the definition of rank $(1, 0)$, means that whenever such trees exist then we can attempt to use square function estimates on our collection of P_j tiles that come with those trees. In other words, the P_j tiles stack up similarly as in the bilinear Hilbert transform case.

Recall the John-Nirenberg inequality [13].

Lemma 9.2. *Let $\vec{\mathbf{P}}$ be a finite collection of quadtiles, $j = 1, 2, 3, 4$ and let $(a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}$ be a sequence of complex numbers. Then*

$$size_j((a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}) \sim \sup_{T \subset \vec{\mathbf{P}}} \frac{1}{|I_T|} \left\| \left(\sum_{\vec{P} \in T} |a_{P_j}|^2 \frac{1_{I_{\vec{P}}}}{|I_{\vec{P}}|} \right)^{1/2} \right\|_{L^{1,\infty}(I_T)}$$

where T ranges over all trees in $\vec{\mathbf{P}}$ which are either one quadtile trees or i-trees for some $2 \leq i \leq 4$ such that j is a good index with respect to i, as in the definition of rank $(1, 0)$.

The proof carries exactly over due to our choice of possible trees in the definition of size.

10. PROOF OF DISCRETE OPERATOR THEOREM

Proposition 10.1. *Let $\vec{\mathbf{P}}$ be a finite collection of quadtiles. Then*

$$|\Lambda_{\vec{\mathbf{P}}}(f_1, f_2, f_3, f_4)| \lesssim size((\langle f_1, \phi_{P_1^{n_1}, 1} \rangle)_{\vec{P} \in \vec{\mathbf{P}}}) \prod_{j=2}^{4} (size((\langle f_j, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}}))^{\theta_j}$$

$$(energy((\langle f_j, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}}))^{1-\theta_j}$$

for any $0 \leq \theta_2, \theta_3, \theta_4 < 1$ with $\theta_2 + \theta_3 + \theta_4 = 1$, with the implicit constant depending on the θ_i.

This proposition will be proven in section 14.

Lemma 10.2. *Let $\vec{\mathbf{P}}$ be a finite collection of quadtiles, $j \in \{1, 2, 3, 4\}$ and E be a set of finite measure. Then for every $|f| \leq 1_E$ one has*

$$size((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}}) \lesssim \log_2(\langle n_j \rangle) \sup_{\vec{P} \in \vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|} \int_E \tilde{\chi}_{I_{P_j^{n_j}}}^M$$

for all $M > 0$, with the implicit constant depending on M.

Lemma 10.2 will be proven in section 12.

Define the shifted dyadic maximal operator M^n as follows [11]

$$M^n f(x) := \sup_{x \in I} \frac{1}{|I|} \int_{\mathbb{R}} |f(y)| \tilde{\chi}_{I^n}(y) dy$$

where the supremum is taken only over dyadic intervals.

Lemma 10.3. *For any $n \in \mathbb{Z}$ the shifted maximal function M^n maps boundedly $L^p(\mathbb{R})$ into $L^p(\mathbb{R})$ with a bound of the type $O(\log_2(\langle n \rangle))$. It also maps boundedly $L^\infty(\mathbb{R})$ into $L^\infty(\mathbb{R})$ and $L^1(\mathbb{R})$ into $L^{1,\infty}(\mathbb{R})$ with a bound of the type $O(\log_2(\langle n \rangle))$.*

Lemma 10.3 will be proven in section 11.

Lemma 10.4. *Let $\vec{\mathbf{P}}$ be a finite collection of quadtiles, $j \in \{2, 3, 4\}$ and $f \in L^2(\mathbb{R})$. Then*

$$energy((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}}) \lesssim (\log_2(\langle n_j \rangle))^2 \|f\|^2$$

Lemma 10.4 will be proven in section 13.

We can now prove theorem 7.1.

Proof. Fix $E_1, E_2, E_3, E_4, \gamma_1, \gamma_2$ and γ_3 as in the hypothesis of theorem 7.1. The goal is to find $E_4' \subseteq E_4$ with $|E_4'| \sim |E_4|$ such that for every $|f_1| \lesssim 1_{E_1}, |f_2| \lesssim 1_{E_2}, |f_3| \lesssim 1_{E_3}$ one has

$$\left| \Lambda_{\vec{\mathbf{P}}}(f_1, f_2, f_3, 1_{E_4'}) \right| \lesssim \left(\prod_{j=1}^3 |\log_2(\langle n_j \rangle)|^4 \right) |E_1|^{\gamma_1} |E_2|^{\gamma_2} |E_3|^{\gamma_3} |E_4|^{\gamma_4}$$

where we recall that γ_4 is defined by $\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 = 1$.

Using the dilation symmetry of T_β, which translates naturally to $\Lambda_{\vec{\mathbf{P}}}$, one can clearly assume wlog that $|E_4| = 1$. Define then the set Ω by

$$\Omega := \bigcup_{j=1}^3 \left(\left\{ x : M^{n_j} \left(\frac{1_{E_j}}{|E_j|} \right)(x) > C \log_2(\langle n_j \rangle) \right\} \right)$$

and observe that $|\Omega| \ll 1$ if C is a large enough constant. Then set $E_4' := E_4 \setminus \Omega$ and notice that $|E_4'| \sim 1$ as desired.

Then for any $d \geq 1$ define the collection $\vec{\mathbf{P}}_d$ by

$$\vec{\mathbf{P}}_d := \{\vec{P} \in \vec{\mathbf{P}} : 2^{d-1} \leq \frac{dist(I_{\vec{P}}, \Omega^c)}{|I_{\vec{P}}|} \leq 2^d\}$$

and let P_0 be the collection of quadtiles which intersect Ω^c. Clearly $\bigcup_{d \geq 0} \vec{\mathbf{P}}_d = \vec{\mathbf{P}}$.

We can write

$$(10.1) \qquad \Lambda_{\vec{\mathbf{P}}}(f_1, f_2, f_3, 1_{E_4'}) = \sum_{d=0}^{\infty} \int_{\mathbb{R}} T_{\vec{\mathbf{P}}_d}(f_1, f_2, f_3)(x) 1_{E_4'}(x) dx$$

Fix $d \geq 0$ and consider the inner quad linear form of (10.1). It can be estimated by proposition 10.1. Using lemma 10.2 and lemma 10.3 we obtain

$$\text{size}((\langle f, \phi_{P_j^{n_j},j}\rangle)_{\vec{P} \in \vec{\mathbf{P}}_d}) \lesssim \log_2(\langle n_j \rangle) \sup_{\vec{P} \in \vec{\mathbf{P}}_d} \frac{1}{|I_{\vec{P}}|} \int_E \tilde{\chi}^M_{I_{P_j^{n_j}}}$$

$$\lesssim (\log_2(\langle n_j \rangle))^2 \min(1, 2^d |E_j|)$$

$$\lesssim (\log_2(\langle n_j \rangle))^2 2^d |E_j|^{a_j}$$

for any $0 < a_j < 1$, $j = 1, 2, 3$.

Using lemma 10.4 we also obtain for $j = 2, 3$

$$\text{energy}((\langle f, \phi_{P_j^{n_j},j}\rangle)_{\vec{P} \in \vec{\mathbf{P}}}) \lesssim (\log_2(\langle n_j \rangle))^2 |E_j|^{1/2}.$$

Using lemmas 10.2, 10.4 and 10.3 for the fourth position, using $n_4 = 0$, we note that since $|E_4| = 1$ we obtain

$$\text{size}((\langle f, \phi_{P_j^{n_j},j}\rangle)_{\vec{P} \in \vec{\mathbf{P}}_d}) \lesssim 2^{-Md}$$

and

$$\text{energy}((\langle f, \phi_{P_j^{n_j},j}\rangle)_{\vec{P} \in \vec{\mathbf{P}}}) \lesssim 1.$$

Putting all this together then proposition 10.1 allows us to bound the corresponding quad linear form in (10.1) for a fixed $d \geq 0$ by

$$2^{-\#d}|E_1|^{a_1}(|E_2|^{a_2})^{\theta_2}(|E_2|^{1/2})^{1-\theta_2}(|E_3|^{a_3})^{\theta_2}(|E_3|^{1/2})^{1-\theta_3} \cdot 1$$

$$= 2^{-\#d}|E_1|^{a_1}|E_2|^{a_2\theta_2 + \frac{1}{2}(1-\theta_2)}|E_3|^{a_3\theta_3 + \frac{1}{2}(1-\theta_3)}$$

where $\#$ is a strictly positive integer. Then we can make a_1 arbitrarily close to 1, $a_2\theta_2 + \frac{1}{2}(1 - \theta_2)$ arbitrarily close to $\frac{1}{2}$ by choosing θ_2 close to 0 and $a_3\theta_3 + \frac{1}{2}(1 - \theta_3)$ arbitrarily close to 1 by choosing θ_3 close to 1 and a_3 also close to 1. $\qquad \square$

11. ESTIMATES FOR THE SHIFTED DYADIC MAXIMAL FUNCTION

We will now recall the proof of lemma 10.3 from [11]. We note, as Muscalu does in [11], that the proof of this lemma was already known and can be found in [16] Chapter II.

Proof. Observe that it is sufficient to prove the estimates for the "sharp" shifted dyadic maximal function $\tilde{M}^n$ defined by

$$\tilde{M}^n f(x) := \sup_{x \in I} \frac{1}{|I|} \int_{I^n} |f(y)| dy$$

where the supremum is only taken over dyadic intervals.

To observe this, fix x and I so that $x \in I$. We can write

$$\frac{1}{|I^n|} \int_{I^n} |f(y)| dy \lesssim \sum_{\# \in \mathbb{Z}} \left[\frac{1}{|I^{n+\#}|} \int_{I^{n+\#}} |f(y)| dy \right] \frac{1}{\langle \# \rangle^{100}}.$$

Assuming the theorem holds for $\tilde{M}^n$ and using the above, one has

$$
\begin{aligned}
\|M^n f\|_p &\lesssim \sum_{\# \in \mathbb{Z}} \frac{1}{\langle \# \rangle^{100}} \|\tilde{M}^{n+\#} f\|_p \\
&\lesssim \sum_{\# \in \mathbb{Z}} \frac{1}{\langle \# \rangle^{100}} \log_2(\langle n + \# \rangle) \|f\|_p \\
&\lesssim \sum_{\# \in \mathbb{Z}} \frac{1}{\langle \# \rangle^{100}} \log_2(\langle n \rangle \langle \# \rangle) \|f\|_p \\
&\lesssim \log_2(\langle n \rangle) \|f\|_p
\end{aligned}
$$

as desired. We then turn to proving the theorem for $\tilde{M}^n$.

Let $\lambda > 0$. We claim that the following inequality is true

(11.1) $$|\{x : \tilde{M}^n f(x) > \lambda\}| \lesssim \log_2(\langle n \rangle) |\{x : Mf(x) > \lambda\}|$$

where M is the classical Hardy-Littlewood maximal operator. Assuming (11.1) the theorem for $\tilde{M}^n$ follows from the Hardy-Littlewood theorem in the case $L^1(\mathbb{R}) \mapsto L^{1,\infty}(\mathbb{R})$. The case $L^\infty(\mathbb{R}) \mapsto L^\infty(\mathbb{R})$ is trivial. All the other estimates we obtain then by interpolating between those two cases.

To prove (11.1) denote by $\mathcal{I}_n^\lambda$ the collection of all dyadic and maximal, with respect to inclusion, intervals I^n, for which

$$
\frac{1}{|I^n|} \int_{I^n} |f(y)| dy > \lambda.
$$

Observe they are all disjoint and in addition one has

$$
\bigcup_{I^n \in \mathcal{I}_n^\lambda} I^n = \{x : Mf(x) > \lambda\}.
$$

For every such selected, maximal, dyadic interval I^n, then it has at most $\log_2(\langle n \rangle)$ friends as in the tree case. More precisely then there are at most $\log_2(\langle n \rangle)$ disjoint dyadic intervals $I_1^n, \ldots, I_N^n$ of the same length as $|I^n|$, so that the translate with $-n$ corresponding units of any subinterval of I^n becomes a subinterval of one of these intervals. Now we claim

$$
\{x : \tilde{M}^n f(x) > \lambda\} \subseteq \bigcup_{I^n \in \mathcal{I}_n^\lambda} (I_1^n \cup \ldots \cup I_1^N).
$$

To prove this, pick x^* such that $M^n f(x^*) > \lambda$. This implies that there exists a dyadic interval J containing x^* such that $\frac{1}{|J^n|} \int_{J^n} |f(y)| dy > \lambda$. Due to the previous construction, one can certainly find one selected maximal interval of the type I^n such that $J^n \subseteq I^n$. This however means that J itself will be a subset of one of $I_1^n, \ldots, I_N^n$ which proves the claim.

One can now easily see that this claim and the disjointness of the maximal intervals I^n along with the fact that $N \leq \log_2(\langle n \rangle)$ imply (11.1).

$\square$

12. Size estimates

We will now prove lemma 10.2.

Proof. Fix $j \in \{1, 2, 3, 4\}$, n_j, E and $|f| \lesssim 1_E$ as in the lemma. Since $\vec{\mathbf{P}}$ is a finite set of tiles there exists a tree $\tilde{T}$ such that the supremum in the size is attained. If the tree is just one quadtile then the proof is trivial. Let's thus assume that $\tilde{T}$ is an i-tree for some $2 \le i \le 4$ such that j is a good index with respect to i, as in the definition of rank $(1, 0)$.

$$(12.1) \quad \text{size}((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}}) = (\frac{1}{|I_{\tilde{T}}|} \sum_{\vec{P} \in \tilde{T}} |\langle f, \phi_{P_j^{n_j}, j} \rangle|^2)^{1/2}$$

$$\le \sum_{i \in Fr(n_j)} (\frac{1}{|I_{\tilde{T}}|} \sum_{\substack{\vec{P} \in \tilde{T} \\ I_{\vec{P}} \subseteq I_{\tilde{T}}^i}} |\langle f, \phi_{P_j^{n_j}, j} \rangle|^2)^{1/2}$$

Now for each $i \in Fr(n_j)$ take $\vec{P} \in \tilde{T}$ such that $I_{\vec{P}} \subseteq I_{\tilde{T}}^i$ and pick from that collection of tiles trees that are maximal with regards to inclusion and such that they contain their top. Call that collection $\vec{T}_i$ for each $i \in Fr(n_j)$. Then we can bound (12.1) with

$$\sum_{i \in Fr(n_j)} \sum_{T \in \vec{T}_i} (\frac{1}{|I_{\tilde{T}}|} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2)^{1/2}$$

Note that the trees in $\vec{T}_i$ are disjoint and in particular

$$\sum_{T \in \vec{T}_i} |I_T| \le |I_{\tilde{T}}|.$$

Thus for a fixed $i \in Fr(n_j)$ we have

$$\sum_{T \in \vec{T}_i} (\frac{1}{|I_{\tilde{T}}|} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2)^{1/2} \le \left(\sup_{T \in \vec{T}_i} \left(\frac{1}{|I_T|} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2 \right)^{1/2} \right) \frac{1}{|I_{\tilde{T}}|} \sum_{T \in \vec{T}_i} |I_T|$$

$$\le \sup_{T \in \vec{T}_i} \left(\frac{1}{|I_T|} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2 \right)^{1/2}$$

Since $\vec{\mathbf{P}}$ is a finite set of tiles then for each friend there exists a tree T which is an i-tree for some $i \ne j$, $2 \le i \le 4$, such that

$$\sup_{T \in \vec{T}_i} \left(\frac{1}{|I_T|} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2 \right)^{1/2} \sim \frac{1}{|I_T|} \left\| \left(\sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2 \frac{1_{I_{\vec{P}}}}{|I_{\vec{P}}|} \right)^{1/2} \right\|_{1, \infty}$$

Here we have also used the John-Nirenberg inequality in lemma 9.2. Clearly it is enough to prove that

$$\left\| \left(\sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2 \frac{1_{I_{\vec{P}}}}{|I_{\vec{P}}|} \right)^{1/2} \right\|_{1, \infty} \lesssim \int_{\mathbb{R}} 1_{E_j} \tilde{\chi}_{I_T}^M$$

and use the fact that $|Fr(n_j)| \le \log_2(\langle n_j \rangle)$.

Decompose the real line as a union of intervals

$$\mathbb{R} = \bigcup_{n \in \mathbb{Z}} I_T^n$$

where $|I_T^n| = |I_T|$ for every $n \in \mathbb{Z}$, $I_T^0 = I_T$ and all I_T^n are disjoin except for the endpoints. We think of I_T^n as being n units of length $|I_T|$ to the right of I_T if $n > 0$ and to the left if $n < 0$. Then split f as

$$f = f \cdot 1_{5I_T} + f \cdot 1_{(5I_T)^c}.$$

Since the expression $\left(\sum_{\vec{P} \in T} |\langle f, \phi_{P_j, j} \rangle|^2 \frac{1_{I_{\vec{P}}}}{|I_{\vec{P}}|} \right)^{1/2}$ is a square function, it is bounded from L^1 into $L^{1,\infty}$ and as a consequence

$$\left\| \left(\sum_{\vec{P} \in T} |\langle f \cdot 1_{5I_T}, \phi_{P_j, j} \rangle|^2 \frac{1_{I_{\vec{P}}}}{|I_{\vec{P}}|} \right)^{1/2} \right\|_{1,\infty} \lesssim \|f \cdot 1_{5I_T}\|_1$$

which can be majorized by the expression in the right-hand side of the lemma.

We are left with estimating

$$\left\| \left(\sum_{\vec{P} \in T} |\langle f \cdot 1_{(5I_T)^c}, \phi_{P_j, j} \rangle|^2 \frac{1_{I_{\vec{P}}}}{|I_{\vec{P}}|} \right)^{1/2} \right\|_{1,\infty}$$

which is clearly smaller than

$$\sum_{|n| \geq 3} \sum_{\vec{P} \in T} \frac{\langle |f \cdot 1_{I_T^n}|, |\phi_{P_j, j}| \rangle}{|I_P|^{1/2}} |I_P| \lesssim \sum_{|n| \geq 3} \sum_{\vec{P} \in T} \langle |f| \cdot 1_{I_T^n}, |\tilde{\chi}_{I_{\vec{P}}}^M| \rangle$$

for any big number $M > 0$. In order to complete the proof it is enough to prove that

$$\sum_{\vec{P} \in T} \langle |f| \cdot 1_{I_T^n}, |\tilde{\chi}_{I_{\vec{P}}}^M| \rangle \lesssim \frac{1}{\langle n \rangle^M} \int_{\mathbb{R}} 1_{E_j} 1_{I_T^n}$$

but this is an easy consequence of the fact that the sum on the left-hand side runs over P for which $I_P \subseteq I_T$. This ends the proof of lemma 10.2.

$\square$

13. Energy estimates

We will now prove lemma 10.4.

Proof. Fix $j \in \{1, 2, 3, 4\}$ and $f \in L^2(\mathbb{R})$. Let also n and $\mathbf{T}$ be as in definition of energy such that the supremum in the definition is attained. We want to show that

$$(13.1) \qquad 2^n \left(\sum_{T \in \mathbf{T}} |I_T| \right)^{1/2} \lesssim \|f\|_2$$

If we square the left-hand side of (13.1) and use the properties of the trees in $\mathbf{T}$ we can write

$$\left(2^n \left(\sum_{T \in \mathbf{T}} |I_T|\right)^{1/2}\right)^2 = 2^{2n} \sum_{T \in \mathbf{T}} |I_T|$$

$$\lesssim 2^{2n} 2^{-2n} \sum_{T \in \mathbf{T}} \left(\sum_{\vec{P} \in T} |\langle f, \phi_{P_j^{n_j}, j}\rangle|^2\right) = \sum_{T \in \mathbf{T}} \left(\sum_{\vec{P} \in T} |\langle f, \phi_{P_j^{n_j}, j}\rangle|^2\right)$$

and this expression is supposed to be smaller than $\|f\|_2^2$. We can also write

$$\sum_{T \in \mathbf{T}} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j^{n_j}, j}\rangle|^2 = |\langle \sum_{T \in \mathbf{T}} \sum_{\vec{P} \in T} \langle f, \phi_{P_j^{n_j}, j}\rangle \phi_{P_j^{n_j}, j}, f\rangle|$$

$$\lesssim \|f\|_2 \| \sum_{T \in \mathbf{T}} \sum_{\vec{P} \in T} \langle f, \phi_{P_j^{n_j}, j}\rangle \phi_{P_j^{n_j}, j}\|_2$$

so it is enough to prove that

(13.2)
$$\| \sum_{T \in \mathbf{T}} \sum_{\vec{P} \in T} \langle f, \phi_{P_j^{n_j}, j}\rangle \phi_{P_j^{n_j}, j}\|_2 \lesssim \left(\sum_{T \in \mathbf{T}} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j^{n_j}, j}\rangle|^2\right)^{1/2}$$

The square of the left-hand side of (13.2) becomes smaller than

(13.3)
$$\sum_{T, T' \in \mathbf{T}} \sum_{\substack{\vec{P} \in T \\ \vec{Q} \in T'}} |\langle f, \phi_{P_j^{n_j}, j}\rangle| |\langle f, \phi_{Q_j^{n_j}, j}\rangle| |\langle \phi_{P_j^{n_j}, j}, \phi_{Q_j^{n_j}, j}\rangle| := \mathrm{I} + \mathrm{II}$$

where I contains the part where $T \neq T'$ while II contains the $T = T'$ part.

We first estimate I. Observe that if $\vec{P} \in T$ and $\vec{Q} \in T'$ then, in order for $\langle \phi_{P_j^{n_j}, j}, \phi_{Q_j^{n_j}, j}\rangle$ to be non-zero, we must have $\omega_{P_j} \cap \omega_{Q_j} \neq \emptyset$ and so we either have $\omega_{P_j} \subseteq \omega_{Q_j}$ or $\omega_{Q_j} \subseteq \omega_{P_j}$. Because of the symmetry we can assume that we always have $\omega_{P_j} \subseteq \omega_{Q_j}$. Then, since T and T' are strictly disjoint, this means that $I_{\vec{Q}} \cap I_T = \emptyset$ for any such a $\vec{Q}$.

Fix now $T, T', \vec{P} \in T$ and $\vec{Q} \in T'$ so that $\omega_{P_j} \subseteq \omega_{Q_j}$. Using the properties of the trees $T \in \mathbf{T}$, we can write

$$\frac{1}{|I_{\vec{P}}|^{1/2}} |\langle f, \phi_{P_j^{n_j}, j}\rangle| \lesssim 2^n \lesssim \frac{1}{|I_T|^{1/2}} \left(\sum_{\tilde{P}} |\langle f, \phi_{\tilde{P}_j^{n_j}, j}\rangle|^2\right)^{1/2}$$

from which we can deduce that

(13.4)
$$|\langle f, \phi_{P_j^{n_j}, j}\rangle| \lesssim \frac{|I_{\vec{P}}|^{1/2}}{|I_T|^{1/2}} \left(\sum_{\tilde{P}} |\langle f, \phi_{\tilde{P}_j^{n_j}, j}\rangle|^2\right)^{1/2}.$$

Similarly we have

(13.5)
$$|\langle f, \phi_{Q_j^{n_j}, j}\rangle| \lesssim \frac{|I_{\vec{Q}}|^{1/2}}{|I_T|^{1/2}} \left(\sum_{\tilde{P}} |\langle f, \phi_{\tilde{P}_j^{n_j}, j}\rangle|^2\right)^{1/2}.$$

Using (13.4) and (13.5) we can bound I in (13.3) with

$$(13.6) \quad \sum_{T,T'\in\mathbf{T}} \sum_{\substack{\vec{P}\in T \\ \vec{Q}\in T' \\ \omega_{P_j}\subseteq\omega_{Q_j}}} \left[\frac{|I_{\vec{P}}|^{1/2}}{|I_T|^{1/2}}\left(\sum_{\vec{P}}|\langle f,\phi_{\tilde{P}_j^{n_j},j}\rangle|^2\right)^{1/2}\right]\left[\frac{|I_{\vec{Q}}|^{1/2}}{|I_T|^{1/2}}\left(\sum_{\vec{P}}|\langle f,\phi_{\tilde{P}_j^{n_j},j}\rangle|^2\right)^{1/2}\right]$$

$$|\langle\phi_{P_j^{n_j},j},\phi_{Q_j^{n_j},j}\rangle|$$

$$= \sum_{T\in\mathbf{T}}\left(\sum_{\vec{P}\in T}|\langle f,\phi_{\tilde{P}_j^{n_j},j}\rangle|^2\right)\sum_{\vec{P}\in T}\sum_{\substack{T'\in\mathbf{T}\\T'\neq T}}\sum_{\substack{\vec{Q}\in T'\\\omega_{P_j}\subseteq\omega_{Q_j}}}\frac{1}{|I_T|}|I_{\vec{P}}|^{1/2}|I_{\vec{Q}}|^{1/2}|\langle\phi_{P_j^{n_j},j},\phi_{Q_j^{n_j},j}\rangle|$$

$$\lesssim \sum_{T\in\mathbf{T}}\left(\sum_{\vec{P}\in T}|\langle f,\phi_{\tilde{P}_j^{n_j},j}\rangle|^2\right)\sum_{\vec{P}\in T}\sum_{\substack{T'\in\mathbf{T}\\T'\neq T}}\sum_{\substack{\vec{Q}\in T'\\\omega_{P_j}\subseteq\omega_{Q_j}}}\frac{1}{|I_T|}|\langle\tilde{\chi}_{I_{P_j^{n_j}}},\tilde{\chi}_{I_{Q_j^{n_j}}}\rangle|$$

Fix T and look at the corresponding inner sum in (13.6).

$$(13.7) \quad \sum_{\vec{P}\in T}\sum_{\substack{T'\in\mathbf{T}\\T'\neq T}}\sum_{\substack{\vec{Q}\in T'\\\omega_{P_j}\subseteq\omega_{Q_j}}}\frac{1}{|I_T|}|\langle\tilde{\chi}_{I_{P_j^{n_j}}},\tilde{\chi}_{I_{Q_j^{n_j}}}\rangle|$$

It is clearly enough to show that this expression is $O((\log_2(\langle n_j\rangle))^2|I_T|)$.

Fix $\vec{P}\in T$ and recall

$$|\langle\tilde{\chi}_{I_{P_j^{n_j}}},\tilde{\chi}_{I_{Q_j^{n_j}}}\rangle|\lesssim\left(1+\frac{\text{dist}(I_{P_j^{n_j}},I_{Q_j^{n_j}})}{|I_{\vec{P}}|}\right)^{-M}|I_{\vec{Q}}|.$$

Set $\vec{\mathbf{Q}}_{\vec{P}}=\{\vec{Q}\in T':T'\in\mathbf{T},T'\neq T,\omega_{P_j}\subseteq\omega_{Q_j}\}$. Pick $\vec{\tilde{Q}}$ from $\vec{\mathbf{Q}}_{\vec{P}}$ such that $I_{\tilde{Q}_j^{n_j}}$ is maximal with respect to inclusion and place all $\vec{\tilde{\tilde{Q}}}\in\vec{\mathbf{Q}}_{\vec{P}}$ such that $I_{\tilde{\tilde{Q}}_j^{n_j}}\cap I_{\tilde{Q}_j^{n_j}}\neq\emptyset$ and $\vec{\tilde{\tilde{Q}}}\neq\vec{\tilde{Q}}$ into $S_{\vec{\tilde{Q}}}$. Then observe that

$$\sum_{\vec{Q}\in S_{\vec{\tilde{Q}}}\cup\{\tilde{Q}\}}|\langle\tilde{\chi}_{I_{P_j^{n_j}}},\tilde{\chi}_{I_{Q_j^{n_j}}}\rangle|\lesssim\sum_{\vec{Q}\in S_{\vec{\tilde{Q}}}\cup\{\tilde{Q}\}}\left(1+\frac{\text{dist}(I_{P_j^{n_j}},I_{Q_j^{n_j}})}{|I_{\vec{P}}|}\right)^{-M}|I_{\vec{Q}}|$$

$$\lesssim\left(1+\frac{\text{dist}(I_{P_j^{n_j}},I_{\tilde{Q}_j^{n_j}})}{|I_{\vec{P}}|}\right)^{-M}\sum_{\vec{Q}\in S_{\vec{\tilde{Q}}}\cup\{\tilde{Q}\}}|I_{\vec{Q}}|.$$

Here we use the fact that $|I_{\vec{P}}|>|I_{\vec{Q}}|$ for all $\vec{Q}\in\vec{\mathbf{Q}}_{\vec{P}}$. Now note that the $I_{\vec{Q}}$ for all $\vec{Q}\in S_{\vec{\tilde{Q}}}$ are disjoint and they can only come from the friends of $I_{\vec{\tilde{Q}}}$ so

$$\sum_{\vec{Q}\in S_{\vec{\tilde{Q}}}\cup\{\tilde{Q}\}}|I_{\vec{Q}}|\lesssim\log_2(\langle n_j\rangle)|I_{\vec{\tilde{Q}}}|$$

Now place $\vec{\tilde{Q}}$ into $\vec{\mathbf{Q}}_{\vec{P}}^*$ and throw away $S_{\vec{\tilde{Q}}} \cup \vec{\tilde{Q}}$ from $\vec{\mathbf{Q}}_{\vec{P}}$ and iterate the selection process. Since $\mathbf{P}$ is finite then our selection process will take finitely many steps. We can bound (13.7) from above with

$$(13.8) \qquad \sum_{\vec{P} \in T} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \log_2(\langle n_j \rangle) \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}|$$

where all the $I_{\vec{Q}}$ for $\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*$ are disjoint.

Now split (13.8) in the following way

$$\log_2(\langle n_j \rangle) \sum_{\substack{\vec{P} \in T \\ 4n_j|I_{\vec{P}}| \geq |I_T|}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}|$$

$$+ \log_2(\langle n_j \rangle) \sum_{\substack{\vec{P} \in T \\ 4n_j|I_{\vec{P}}| < |I_T|}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}|.$$

Pick all $\vec{P} \in T$ with $|I_{\vec{P}}|$ of the same length such that $4n_j|I_{\vec{P}}| \geq |I_T|$. Then for a fixed $\vec{P}$ we can estimate

$$\sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}| \lesssim |I_{\vec{P}}|$$

and since the $I_{\vec{P}}$ are all disjoint for $\vec{P} \in T$ of the same scale then when we add up $|I_{\vec{P}}|$ for all of them we get something less than $|I_T|$. Now note there are at most $O(\log_2(\langle n_j \rangle))$ scales of $\vec{P}$ such that $4n|I_{\vec{P}}| = 2^{\log_2(4n_j)}|I_{\vec{P}}| > |I_T|$ and thus

$$\log_2(\langle n_j \rangle) \sum_{\substack{\vec{P} \in T \\ 4n_j|I_{\vec{P}}| \geq |I_T|}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}| \lesssim (\log_2(\langle n_j \rangle))^2 |I_T|.$$

Now look at $\vec{P} \in T$ with $4n_j|I_{\vec{P}}| < |I_T|$. Those $\vec{P}$, that are less than $3n_j$ units of length $|I_{\vec{P}}|$ away from the endpoints of I_T, might interact with $\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*$ and for those we estimate

$$\sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}| \lesssim |I_{\vec{P}}|.$$

Note that for a given scale there are at most $6n_j$ of them. For those that are $l > 3n_j$ units of length $|I_{\vec{P}}|$ away from the endpoints of I_T then $I_{\vec{P}} \cap I_{\vec{Q}} = \emptyset$ for all $\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*$. Thus we estimate

$$\sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}| \lesssim (1 + (l - 3n))^{-M} |I_{\vec{P}}|.$$

For a given such scale of $\vec{P}$, say $|I_{\vec{P}}| = 2^k$, we get

$$\log_2(\langle n_j \rangle) \sum_{\substack{\vec{P} \in T \\ |I_{\vec{P}}| = 2^k}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}|$$

$$\lesssim (\log_2(\langle n_j \rangle)) \left(6n_j |I_{\vec{P}}| + |I_{\vec{P}}| \sum_{l=3n+1}^{\infty} \frac{1}{(1 + (l - 3n_j))^M} \right)$$

$$\lesssim \log_2(\langle n_j \rangle)(6n_j + 1)|I_{\vec{P}}|$$

Now if we sum up over all scales such that $|I_{\vec{P}}| < \frac{|I_T|}{4n_j}$ we get

$$\log_2(\langle n_j \rangle) \sum_{\substack{\vec{P} \in T \\ 4n_j |I_{\vec{P}}| < |I_T|}} \sum_{\vec{Q} \in \vec{\mathbf{Q}}_{\vec{P}}^*} \left(1 + \frac{\text{dist}(I_{P_j^{n_j}}, I_{Q_j^{n_j}})}{|I_{\vec{P}}|} \right) |I_{\vec{Q}}| \lesssim \log_2(\langle n_j \rangle)(6n_j + 1)\frac{|I_T|}{4n_j}$$

$$\lesssim \log_2(\langle n_j \rangle)|I_T|.$$

We are now left with the diagonal term II from (13.3) where the sum runs over $T = T'$. If $\vec{P}, \vec{Q} \in T$ and $\omega_{P_j} \cap \omega_{Q_j} \neq \emptyset$ then we must have $\omega_{P_j} = \omega_{Q_j}$. We can majorize II with

$$\sum_{T \in \mathbf{T}} \sum_{\vec{P} \in T} |\langle f, \phi_{P_j^{n_j}, j} \rangle|^2 \frac{1}{|I_{\vec{P}}|} \left(\sum_{\substack{\vec{Q} \in T \\ \omega_{P_j} = \omega_{Q_j}}} |\langle \tilde{\chi}_{I_{P_j^{n_j}}}, \tilde{\chi}_{I_{Q_j^{n_j}}} \rangle| \right)$$

and it is sufficient to show that

$$\sum_{\substack{\vec{Q} \in T \\ \omega_{P_j} = \omega_{Q_j}}} |\langle \tilde{\chi}_{I_{P_j^{n_j}}}, \tilde{\chi}_{I_{Q_j^{n_j}}} \rangle|$$

is $O(\log_2(\langle n_j \rangle)|I_{\vec{P}}|)$ but that follows immediately from the fact that all the $I_{\vec{Q}}$ for which $\omega_{P_j} = \omega_{Q_j}$ are disjoint.

This concludes the proof of lemma 10.4.

$\square$

14. Proof of proposition 10.1

We will now prove proposition 10.1. Fix the collection $\vec{\mathbf{P}}$ of quadtiles and the functions f_1, f_2, f_3, f_4. As mentioned before then we assume that $\vec{\mathbf{P}}$ is sparse and of rank $(1, 0)$ and assume it is with respect to $\{\{2, 3, 4\}, \{1\}\}$ without loss of generality.

Denote for simplicity

$$S_j := \text{size}((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}})$$

for $j \in \{1, 2, 3, 4\}$ and

$$E_j := \text{energy}((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}})$$

for $j \in \{2, 3, 4\}$.

Proposition 14.1. *Let $j \in \{2, 3, 4\}$ and $\vec{\mathbf{P}}' \subseteq \vec{\mathbf{P}}$, $n \in \mathbb{Z}$ so that*

$$size((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}'}) \leq 2^{-n} E_j.$$

Then one can decompose $\vec{\mathbf{P}}' = \vec{\mathbf{P}}'' \cup \vec{\mathbf{P}}'''$ such that

$$size((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}''}) \leq 2^{-n-1} E_j$$

and $\vec{\mathbf{P}}'''$ can be written as a disjoint union of trees $T \in \mathbf{T}$ such that

$$\sum_{T \in \mathbf{T}} |I_T| \lesssim 2^{2n}$$

Proof. Our rank $(1, 0)$ collection of quadtiles has all the relevant features in common with the collection of tritiles in the bilinear Hilbert transform so the proof from there works here. $\qquad\square$

By iterating the previous result we obtain the following corollary.

Corollary 14.2. *Let $\vec{\mathbf{P}}$ be a finite collection. Then one can split $\vec{\mathbf{P}}$ as*

$$\vec{\mathbf{P}} = \bigcup_{n \in \mathbb{Z}} \vec{\mathbf{P}}_n$$

where for each $n \in \mathbb{Z}$ and $j = 2, 3, 4$ we have

$$size((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in \vec{\mathbf{P}}_n}) \leq \min(2^{-n} E_j, S_j).$$

Also one can cover $\vec{\mathbf{P}}_n$ by a collection of trees $T \in \mathbf{T}_n$ for which

$$\sum_{T \in \mathbf{T}_n} |I_T| \lesssim 2^{2n}.$$

Lemma 14.3. *Let T be an i-tree, $i = 2, 3$ or 4, in $\vec{\mathbf{P}}$ and f_1, f_2, f_3, f_4 fixed functions, then*

$$\sum_{\vec{P} \in T} \frac{1}{|I_{\vec{P}}|} |\langle f_1, \phi_{P_1^{n_1}, 1} \rangle| |\langle f_2, \phi_{P_2^{n_2}, 2} \rangle| |\langle f_3, \phi_{P_3^{n_3}, 3} \rangle| |\langle f_4, \phi_{P_4^{n_4}, 4} \rangle| \leq |I_T| \prod_{j=1}^{4} size((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in T})$$

Proof. Say T is a 2-tree and assume without loss of generality that 1 and 4 are good indices with respect to the index 2. This is for example the case for our particular operator when we are in the case (4.2) as discussed in section 5. We can bound the left-hand side by

$$\left(\sum_{\vec{P} \in T} |\langle f_1, \phi_{P_1^{n_1}, 1} \rangle| \right)^{1/2} \left(\sup_{\vec{P} \in T} \frac{|\langle f_2, \phi_{P_2^{n_2}, 2} \rangle|}{|I_{\vec{P}}|^{1/2}} \right) \left(\sup_{\vec{P} \in T} \frac{|\langle f_3, \phi_{P_2^{n_3}, 3} \rangle|}{|I_{\vec{P}}|^{1/2}} \right) \left(\sum_{\vec{P} \in T} |\langle f_4, \phi_{P_4^{n_4}, 4} \rangle| \right)^{1/2}$$

Since 1 and 4 are good indices with respect to 2 we clearly have for $j = 1, 4$

$$\left(\sum_{\vec{P} \in T} |\langle f_j, \phi_{P_j^{n_j}, j} \rangle| \right)^{1/2} \leq |I_T|^{1/2} size((\langle f, \phi_{P_j^{n_j}, j} \rangle)_{\vec{P} \in T}).$$

Since trees that consist of a single quadtile are also used in the definition of size then we clearly also have for $j = 2, 3$

$$\sup_{\vec{P}\in T} \frac{|\langle f_j, \phi_{P_j^{n_j},j}\rangle|}{|I_{\vec{P}}|^{1/2}} \le \text{size}((\langle f, \phi_{P_j^{n_j},j}\rangle)_{\vec{P}\in T}).$$

In a similar manner one can verify the lemma for all other possible trees.

$\square$

We now have the tools to complete the proof of proposition 10.1.

Proof. Using the corollary and lemma above then the proof runs as in the bilinear Hilbert transform case. $\square$

15. The Water Wave Problem

In the 2-d water wave problem, Wu showed that if one starts with small initial data then classical solutions exist for a long time [18]. In a natural way she came across operators of the following type

$$f \mapsto p.v. \int_{\mathbb{R}} F\left(\frac{A(x)-A(y)}{x-y}\right) \frac{\Pi_{i=1}^n (B_i(x)-B_i(y))}{(x-y)^{n+1}} f(y) \; dy$$

and had to obtain L^p estimates for them. For such operators L^p estimates are known if $A', B_i' \in L^\infty(\mathbb{R})$ for $i=1,\ldots,n$ and $f \in L^2(\mathbb{R})$. The novelty in Wu's paper was that she faced $B_1' \in L^2(\mathbb{R})$, which indicated that the operator should be viewed as a multilinear operator.

It is clear that operators similar to Wu's appear in PDEs. Just as Calderón commutators appear very naturally in many applications in PDEs and the bilinear Hilbert transform also appears in applications, such as the AKNS systems [14], it is natural to anticipate that operators of a similar type as Wu faces, but with an average dropped, will appear. Thus it is of interest to obtain L^p estimates for operators of the following type

$$(A, b, f) \mapsto p.v. \int_{\mathbb{R}} F\left(\frac{A(x+t)-A(x)}{t}\right) b(x+\beta t) f(x+t)\frac{1}{t} \; dt$$

where F is an analytic function. The first step would be to obtain L^p estimates for

$$(A, b, f) \mapsto p.v. \int_{\mathbb{R}} \left(\frac{A(x+t)-A(x)}{t}\right)^m b(x+\beta t) f(x+t)\frac{1}{t} \; dt$$

with polynomial bounds in m. Theorem 1.1 is the first step in showing a wide range of L^p estimates for such operators when $m=1$.

References

[1] Benyi, A., Demeter, C., Nahmod, A., Thiele, C., Torres, R. and Villarroya, P., *Modulation invariant bilinear T(1) theorem*, Annal. Math. **109** (2009), 279–352.

[2] Calderón, A. P., *Commutators of singular integral operators*, Proc. Nat. Acad. Sci. **53** (1965), 1092–1099.

[3] Christ, M., Journe, J.-L., *Polynomial growth estimates for multilinear singular integral operators*, Acta Math. **159** (1987), 51–80.

[4] Coifman, R. R. and Meyer, Y., *On commutators of singular integrals and bilinear singular integrals*, Trans. Amer. Math. Soc. **212** (1975), 315–331.

[5] Coifman, R. R., McIntosh, A. and Meyer, Y., *L'integrale de Cauchy definit un operateur borne sur L2 pour les courbes Lipschitziennes*, Anal. Math. **116** (1982), 361–387.

[6] Fefferman, C. L., *Pointwise convergence of Fourier series*, Ann. Math. **98** (1973), 551–571.

[7] Grafakos, L. and Li, X., *Uniform bounds for the bilinear Hilbert transforms I*, Ann. Math. **159** (2004), 889–933.

[8] Grafakos, L. and Torres, R. H., *Multilinear Calderón-Zygmund theory*, Adv. in Math. **165** (2002), 124–164.

[9] Lacey, M., Thiele, C., *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$.*, Ann. Math. **146** (1997), pp. 693–724.

[10] Lacey, M., Thiele, C., *On Calderon's conjecture.*, Ann. Math. **149** (1999), pp. 475–496.

[11] Muscalu, C., *Calderón commutators and the Cauchy integral on Lipschitz curves revisited I. First commutator and generalizations*, arXiv:1201.3845v1.

[12] Muscalu, C., Tao, T., Thiele, C., *L^p estimates for the biest I. The Walsh case*, Math. Ann. **329** (2004), 401–426.

[13] Muscalu, C., Tao, T., Thiele, C., *L^p estimates for the biest II. The Fourier case*, Math. Ann. **329** (2004), 427–461.

[14] Muscalu, C., Tao, T., Thiele, C., *Multilinear operators associated to simplexes of arbitrary length*, arXiv:0712.2420v1.

[15] Muscalu, C., Tao, T., Thiele, C., *Multilinear operators given by singular symbols*, J. Amer. Math. Soc. **15** (2002), 469–496.

[16] Stein, E., *Harmonic analysis: real-variable methods, orthogonality and oscillatory integrals*, Princeton University Press, (1993).

[17] Thiele, C., *A uniform estimate*, Ann. Math. **157** (2002), 1–45.

[18] Wu, S., *Almost global wellposedness of the 2-D full water wave problem*, Inventiones Mathematicae. **177** (2009), 45–135.

Department of Mathematics, Cornell University, Ithaca, NY 14850, USA
E-mail address: `eap48@cornell.edu`